NARATIF : Jurnal Ilmiah Nasional Riset Aplikasi dan Teknik Informatika
Vol. 05 No. 02 Desember 2023
P-ISSN: 2656-7377 || E-ISSN: 2714-8467

DETEKSI ANOMALI DAN SERANGAN *LOW RATE DDOS* DALAM LALU LINTAS JARINGAN MENGGUNAKAN *NAIVE BAYES*

Diash Firdaus[1], Fahira[2], Resa Rianti[3]
Program Studi Informatika [1]
Program Studi Teknik Informatika [2, 3]
Institut Teknologi Nasional[1]
Universitas Logistik dan Bisnis Internasional[2,3]
diashfirdaus@gmail.com [1], 1204044@std.ulbi.ac.id[2], 1204053@std.ulbi.ac.id [3]

Abstrak

Low Rate DDoS merupakan serangan jenis *DDoS* yang sulit di deteksi karena memiliki karakteristik paket yang mirip dengan trafik normal, sehingga dibutuhkan algoritma yang memiliki akurasi tinggi serta memiliki latensi yang cukup rendah dalam memberikan keputusan terhadap trafik di jaringan. Metode *Machine Learning* dengan algoritma *Naïve Bayes* digunakan untuk melakukan deteksi terhadap serangan *Low Rate DDoS* karena memiliki akurasi yang cukup baik. Algoritma yang sering digunakan dalam melakukan deteksi serangan *DDoS* seperti SVM, KNN dan Random Forest akan tetapi algoritma tersebut memiliki latensi yang cukup tinggi karena komputasi yang dibutuhkan cukup kompleks. Dengan demikian Algoritma *Machine Learning* yang digunakan adalah algoritma *Naive Bayes* sebagai model prediksi karena memiliki waktu *train*ing yang cepat. Sedangkan *dataset* yang digunakan adalah CICIDS2017. pada tahap *test*ing dilakukan menggunakan 20% dari CICDS2017. Hasil akhir dari penelitian ini adalah teknik deteksi yang efektif dalam mendeteksi anomali dan serangan *Low Rate DDoS*. Hal yang dilakukan adalah merancang strategi pemodelan yang baik untuk mendeteksi anomali dan serangan *Low Rate DDoS* dengan menggunakan pendekatan algoritma *Naive Bayes*. Model *Machine Learning Naive Bayes* dievaluasi dengan menggunakan metrik akurasi, presisi, *recall*, dan f1 score dan dihasilkan model yang dapat memprediksi anomali dan serangan *Low Rate DDoS* pada lalu lintas jaringan dengan baik. Hasil akurasi yang paling tinggi terdapat pada model *GaussianNB* yaitu dengan akurasi 83,45 % dimana telah dibandingkan oleh model BernoulliNB yang tertinggi hanya mendapatkan akurasi 76,21%.
Kata kunci : Deteksi anomali, Serangan *DDoS*, *Naive Bayes*

Abstract

Low Rate DDoS is a type of DDoS attack that is difficult to detect because it has packet characteristics similar to normal traffic, so it requires an algorithm that has high accuracy and has a fairly low latency in making decisions on traffic on the network. Machine Learning method with Naïve Bayes algorithm is used to detect Low Rate DDoS attacks because it has good accuracy. Algorithms that are often used in detecting DDoS attacks such as SVM, KNN and Random Forest but these algorithms have a fairly high latency because the computations required are quite complex. Thus the Machine Learning Algorithm used is the Naive Bayes algorithm as a prediction model because it has a fast training time. While the dataset used is CICIDS2017. at the testing stage is done using 20% of CICDS2017. The final result of this research is an effective detection technique in detecting anomalies and Low Rate DDoS attacks. What is done is to design a good modeling strategy to detect anomalies and Low Rate DDoS attacks using the Naive Bayes algorithm approach. The Machine Learning Naive Bayes model was evaluated using accuracy, precision, recall, and f1 score metrics and produced a model that can predict anomalies and Low Rate DDoS attacks on network traffic well. The highest accuracy results are found in the GaussianNB model with an accuracy of 83.45% which has been compared by the BernoulliNB model which only gets the highest accuracy of 76.21%.
Keywords : Anomaly Detection, DDoS Attack, Naive Bayes.

I. PENDAHULUAN

Pada era revolusi industri 4.0, perkembangan ilmu pengetahuan dan teknologi semakin cepat berkembang. Revolusi industri 4.0 membuat produk yang mengadopsi model *cloud computing*, yaitu cara memberikan berbagai layanan melalui internet[1]. *Cloud computing* menggunakan sumber daya seperti aplikasi, penyimpanan data, server, basis data, jaringan, dan perangkat lunak komputasi yang tersedia di internet sebagai utilitas. Keamanan *cloud computing* juga menimbulkan masalah, terutama dalam jaringan dan akses data, kontrol cloud infrastruktur, dan menerapkan tindakan keamanan yang tidak mudah karena persyaratan keamanan yang berbeda-beda dari pengguna yang berbeda dalam lingkungan komputasi terdistribusi [2]. Anomali dan serangan *Low Rate DDoS* (*Distributed Denial of Service*) merupakan masalah yang sering dihadapi dalam *cloud computing* [3]. Anomali dapat dianggap sebagai data yang tidak "normal" atau outlier dibandingkan dengan data lainnya [3]. Serangan *Low Rate DDoS* adalah jenis serangan penolakan layanan terdistribusi (*DDoS*) yang ditandai oleh tingkat rendah lalu lintas yang dikirim ke server target. Jenis serangan ini sering digunakan untuk mengakali pertahanan jaringan dan firewall yang dirancang untuk memblokir volume tinggi lalu lintas [4]. Deteksi anomali dan serangan *Low Rate DDoS* merupakan hal yang penting untuk dilakukan agar dapat mengidentifikasi aktivitas yang tidak diinginkan pada jaringan dan mengambil tindakan yang tepat [4].

Metode deteksi anomali dan serangan *DDoS* yang sering digunakan adalah dengan menggunakan algoritma *Machine Learning*. *Machine Learning* adalah cabang *artificial intelligence* yang memfokuskan diri pada pembelajaran data

sehingga dapat membentuk sistem yang dapat belajar sendiri tanpa perlu selalu dilatih ulang oleh manusia. Oleh karena itu, menggunakan *Machine Learning* sebagai solusi akan menghasilkan sistem deteksi yang efisien. Salah satu algoritma yang sering digunakan adalah *Naive Bayes*. Algoritma *Naive Bayes* merupakan salah satu algoritma yang sering digunakan untuk klasifikasi data dan dapat bekerja dengan cepat dan efisien. Penelitian ini bertujuan untuk mengevaluasi kemampuan algoritma *Naive Bayes* dalam mendeteksi anomali dan serangan *DDoS* dalam lalu lintas jaringan dan diharapkan dapat membantu dalam deteksi *Low Rate DDoS* dan anomali dengan tingkat akurasi yang tinggi [5].

Hasil akhir dari penelitian ini adalah teknik deteksi yang efektif untuk mendeteksi anomali dan serangan *Low Rate DDoS*. Kontribusi kami dalam penelitian ini adalah merancang strategi pemodelan yang baik menggunakan algoritma *Machine Learning Naive Bayes* sebagai model prediksi untuk mendeteksi anomali dan *Low Rate DDoS*. Algoritma *Naive Bayes* dipilih sebagai model prediksi karena memiliki waktu komputasi yang cepat dan efisien, prinsip yang mudah dipahami, akurasi tinggi, kemampuan untuk menangani data yang tidak terstruktur, dan kemampuan untuk menangani kemungkinan yang tidak seimbang.

II. TINJAUAN PUSTAKA

Pada beberapa tahun terakhir, *Machine Learning* (ML) telah digunakan untuk membuat sistem yang cerdas dengan cara melatih mesin untuk membuat keputusan. Dengan menggunakan *dataset* sebagai input dan *classifier* sebagai metodenya, ML dapat mengidentifikasi data baru yang memiliki kemiripan dengan data yang telah dikenal. Ada banyak algoritma ML yang dapat digunakan untuk membangun model atau *framework* ML, dan setiap algoritma memiliki kelebihan dan kekurangan yang berbeda tergantung pada *dataset* dan fitur yang digunakan.

1. Anomali

Anomali adalah kejadian atau fenomena yang tidak normal atau tidak biasa terjadi. Dalam bidang data mining, anomali dapat diartikan sebagai sebuah kejadian atau observasi yang tidak sesuai dengan pola atau trend yang terjadi pada data. Anomali dapat dianggap sebagai data yang tidak "normal" atau outlier dibandingkan dengan data lainnya. Anomali dapat terjadi karena beberapa alasan, seperti kesalahan data, kejadian yang tidak terduga, atau bahkan adanya kecurangan. Anomali dapat merugikan dalam beberapa kasus, seperti dalam sistem keamanan jaringan yang dapat mengindikasikan adanya aktivitas yang tidak diinginkan seperti serangan *Low Rate DDoS*. Namun, anomali juga dapat menjadi informasi yang bermanfaat jika dapat dianalisis dengan baik, misalnya dalam menemukan pola-pola baru atau trend yang tidak terduga [3].

2. *Low Rate DDoS*

Low Rate DDoS adalah serangan *DDOS* (*Distributed Denial of Service*) dengan tingkat trafik yang rendah. Serangan ini dapat membuatnya lebih sulit untuk dideteksi dan diatasi, karena tingkat trafik yang rendah membuatnya lebih sulit untuk mengidentifikasi sebagai serangan *DDOS*. *Low Rate DDoS* dapat disamar sebagai trafik normal, sehingga menghindari tindakan pencegahan yang dilakukan oleh administrator sistem. Penyerang dapat menggunakan *Low Rate DDoS* untuk menargetkan situs web atau layanan yang terlindungi dengan tingkat pencegahan yang lebih tinggi, karena tingkat trafik yang rendah membuatnya sulit untuk dikenali sebagai serangan *DDOS* oleh sistem pencegahan [4].

3. Naïve Bayes

Gaussian Naive Bayes Classifier, yaitu teknik pembelajaran mesin untuk mengklasifikasikan sekumpulan data berdasarkan fitur-fiturnya. Sebelum digunakan untuk mengklasifikasikan, perlu dilatih terlebih dahulu menggunakan data yang sudah ada yang merupakan milik kelas tertentu. Meskipun asumsinya adalah bahwa setiap fitur tidak memiliki korelasi atau independen satu sama lain, namun pada kebanyakan kasus asumsi ini ternyata tidak terpenuhi. Namun, meskipun ada korelasi atau ketergantungan antara fitur-fiturnya [5]. Pada *Gaussian Naive Bayes*, digunakan distribusi normal atau Gaussian untuk mewakili nilai probabilistik dari setiap fitur untuk setiap kelas. Pertama, mengestimasi distribusi normal dari data *train*ing dengan menghitung mean (μ) dan variance ($\sigma 2$) untuk setiap fitur, kemudian prior probability ($p(Cn)$) dari setiap kelas [5].

Cara kerja algoritma ini adalah dengan menghitung probabilitas a priori dari setiap kelas (jika digunakan untuk klasifikasi multi-kelas), kemudian menghitung probabilitas likelihood dari setiap fitur pada data yang digunakan untuk melatih model dengan mengasumsikan distribusi normal, dan akhirnya mengkombinasikan kedua probabilitas tersebut menggunakan teorema Bayes untuk menentukan kelas yang paling mungkin untuk data baru yang akan diklasifikasikan. Dalam python, GaussianNB dapat digunakan dengan mengimport library scikit-learn dan menggunakan kelas GaussianNB untuk melatih model dan melakukan klasifikasi [6].

4. Pengukuran Performa

Pengukuran performa pada penelitian ini akan menggunakan *confusion matrix* seperti yang disajikan pada tabel 1. *Confusion matrix* merupakan salah satu metode yang umum digunakan untuk mengevaluasi performa dari model *Machine Learning*, termasuk model yang dibangun dengan menggunakan algoritma *decision tree*. *Confusion matrix* merupakan tabel yang menggambarkan hasil prediksi dari model dibandingkan dengan hasil yang sebenarnya [7].

Untuk menggunakan *confusion matrix*, kita perlu menentukan dua kelas yang akan diprediksi oleh model, biasanya disebut sebagai "positif" dan "negatif". Kemudian, kita akan menghitung jumlah *true positive* (TP), *false positive* (FP), *true negative* (TN), dan *false negative* (FN). Jumlah-jumlah tersebut akan digunakan untuk menghitung beberapa metrik yang biasa digunakan dalam mengevaluasi performa model, seperti akurasi, presisi, dan *recall*.

a. Akurasi adalah persentase dari prediksi yang benar dari seluruh prediksi yang dilakukan model. Akurasi dapat dihitung dengan menggunakan rumus :

$$Accuracy = (TP + TN) / (TP + TN + FP + FN).$$

b. Presisi adalah persentase dari prediksi positif yang benar dari seluruh pred okasi positif yang dilakukan model. Presisi dapat dihitung dengan menggunakan rumus :

$$Precission = TP / (TP + FP).$$

c. *Recall* adalah persentase dari prediksi positif yang benar dari seluruh data positif yang ada pada *dataset*. *Recall* dapat dihitung dengan menggunakan rumus :

$$Recall = TP / (TP + FN).$$

d. *F1 score* merupakan rata-rata harmonis dari presisi dan recall. F1 score dapat dihitung dengan menggunakan rumus :

$$F1\ Score = 2 * (presisi * recall) / (presisi + recall).$$

TABEL I
CONFUSION MATRIX

	Hasil Serangan	**Hasil Normal**
Prediksi Serangan	TP (True Positive)	FP (False Positive)
Prediksi Normal	FN (False Negative)	TN (True Negative)

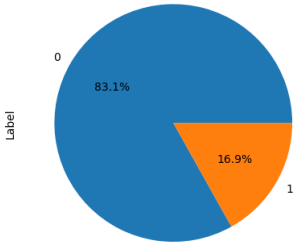
5. *Dataset*

Dataset CICIDS2017 digunakan dalam penelitian. Digunakan pada konferensi CICIS 2017, merupakan kumpulan data *DDoS Low Rate* dari jaringan tradisional. Terdiri dari tipe serangan *DDoS* umum seperti *HTTP Flood*, *UDP Flood*, *SYN Flood*, dll. [8].

Dataset CICIDS2017 digunakan untuk deteksi Anomali dan *DDoS* pada jaringan. Juga digunakan untuk melatih model ML, seperti naive bayes, untuk prediksi Anomali dan serangan *DDoS*. *Dataset* yang digunakan dalam penelitian ini adalah :

TABEL III
DATA PENELITIAN

Jenis Data	**Jumlah**
Serangan	425878
Normal	2096484
Total	2522362

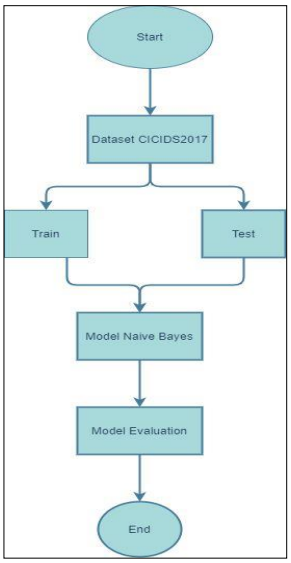


Gambar 1. *Pie Chart* Data Penelitian

III. METODE PENELITIAN

Pada penelitian ini, penulis menggunakan metode penelitian yang telah ditentukan, termasuk menjelaskan kerangka penelitian dan *experiment setup* yang digunakan pada penelitian.

1. Kerangka Penelitian



Gambar 2. Tahapan Penelitian

Penelitian ini, penulis mengikuti beberapa langkah kegiatan yang tertera dalam kerangka penelitian yang ditunjukkan pada gambar 1. Berdasarkan gambar 1, tahap-tahap yang dilakukan dalam penelitian ini adalah sebagai berikut:

a. Penelitian ini, tahap pengumpulan data meliputi penggunaan *dataset* dari CICIDS2017.
b. Setelah selesai mengumpulkan data, *dataset* dari CICIDS2017 akan dibagi menjadi dua bagian, yaitu 80% untuk pelatihan (*train*) dan 20% untuk pengujian (*test*).
c. Kemudian, *dataset train* akan dilatih menggunakan algoritma naive bayes (NB). Setelah proses pelatihan selesai, model yang telah terbentuk akan diuji kemampuannya dengan menggunakan *dataset test*.
d. Evaluasi model, beberapa metrik yang biasa digunakan untuk mengevaluasi hasil pengujian adalah akurasi, presisi, recall, dan f1 score. Akurasi mengukur seberapa baik model mampu memprediksi hasil yang sebenarnya, presisi mengukur seberapa baik model mampu memprediksi hasil positif yang sebenarnya, recall mengukur seberapa baik model mampu menemukan semua hasil positif yang sebenarnya, dan f1 score merupakan rata-rata harmonis dari presisi dan recall. Dengan menghitung metrik-metrik tersebut, kita dapat mengetahui seberapa baik model yang telah dibuat dapat digunakan untuk memprediksi anomali dan serangan *Low Rate DDoS* pada lalu lintas jaringan.

2. Lingkungan Eksperimen

Penelitian ini dilakukan dengan simulasi menggunakan beberapa tools dan system operasi yang bisa dilihat pada Tabel 3

TABEL IIIII
EXPERIMENT SETUP

1	**System OS Virtual**	**Ubuntu dan Kali Linux**
2	Simulation Tools	Mininet, VirtualBox, tcpdump
3	*Dataset*	CICIDS2017
4	CPU	AMD Athlon 300U with Radeon Vega Mobile Gfx
5	RAM	12GB
6	Operating System	Windows 11
7	Supporting Tools	Google Colab Sklearn library Matplotlib Library Python3

IV. HASIL DAN PEMBAHASAN

Pada hasil dan pembahasan merupakan hasil dari model *naïve bayes* untuk mendeteksi anomali dan serangan *Low Rate DDoS* dalam lalu lintas jaringan. Berikut tahapan yang dilakukan dalam penelitian :

1. Fitur & Data

	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	Active Max	Active Min	Idle Mean	Idle Std	Idle Max	Idle Min	Label
0	54865	3	2	0	12	0	6	6	6.0	0.0	...	20	0.0	0.0	0	0	0.0	0.0	0	0	BENIGN
1	55054	109	1	1	6	6	6	6	6.0	0.0	...	20	0.0	0.0	0	0	0.0	0.0	0	0	BENIGN
2	55055	52	1	1	6	6	6	6	6.0	0.0	...	20	0.0	0.0	0	0	0.0	0.0	0	0	BENIGN
3	46236	34	1	1	6	6	6	6	6.0	0.0	...	20	0.0	0.0	0	0	0.0	0.0	0	0	BENIGN
4	54863	3	2	0	12	0	6	6	6.0	0.0	...	20	0.0	0.0	0	0	0.0	0.0	0	0	BENIGN

Gambar 1. Fitur & Data pada *Dataset* CICDS2017

Tahap pertama yang dilakukan adalah menggabungkan semua data yang terdapat pada CICDS2017 dan menampilkan data seperti yang terdapat pada gambar 2 yang menampilkan 5 data teratas pada *dataset*. Setelah, menampilkan fitur dan isi data tahap selanjutnya melakukan pre-processing data.

2. *Pre-processing* data

Pada tahap *pre–processing* merupakan pengelolaan agar data yang didapat menjadi lebih efektif untuk dilakukan *modeling*. Berikut tahapan pre – processing yang telah dilakukan :

a. Melakukan proses encoding

```
0    2096484
1     425878
Name: Label, dtype: int64
```

Gambar 2. Hasil setelah melakukan label encoding

Pada kolom label dilakukan proses encoding pengubahan text menjadi angka agar dipahami python. Seperti pada gambar 3 yang setiap data normal menjadi 0 dan lainnya adalah serangan menjadi 1.

b. Melakukan pemilihan fitur

Tabel 4 merupakan hasil *Feature Selection. Feature Selection* dilakukan untuk meningkatkan akurasi dan mengurangi latensi dikarenakan fitur yang dinilai tidak relevan.

TABEL IVV
FEATURE SELECTION

No.	Features	No.	Features	No.	Features
1	Destination Port	15	Fwd IAT Total	29	Packet Length Std
2	Flow Duration	16	Fwd IAT Mean	30	Packet Length Variance
3	Total Length of Bwd Packets	17	Fwd IAT Std	31	Average Packet Size
4	Fwd Packet Length Mean	18	Fwd IAT Max	32	Avg Fwd Segment Size
5	Fwd Packet Length Std	19	Fwd IAT Min	33	Avg Bwd Segment Size
6	Bwd Packet Length Max	20	Bwd IAT Total	34	Subflow Bwd Bytes
7	Bwd Packet Length Mean	21	Bwd IAT Mean	35	Active Mean
8	Bwd Packet Length Std	22	Bwd IAT Std	36	Active Std
9	Flow Bytes/s	23	Bwd IAT Max	37	Active Max
10	Flow Packets/s	24	Bwd IAT Min	38	Idle Mean
11	Flow IAT Mean	25	Fwd Packets/s	39	Idle Max
12	Flow IAT Std	26	Bwd Packets/s	40	Idle Min
13	Flow IAT Max	27	Max Packet Length	41	Label
14	Flow IAT Min	28	Packet Length Mean		

Selanjutnya Tabel 4 menjelaskan bahwa terdapat fitur dan label yang didapat setelah melalui proses pemilihan fitur dari 78 atribut menjadi 40 fitur dan 1 Label. Hal ini dilakukan dengan menggunakan metode Chi-square (chi2) adalah sebuah uji statistik yang digunakan untuk menguji hipotesis nol. Hipotesis nol adalah asumsi awal bahwa tidak ada perbedaan atau korelasi antara dua variabel. dimana ada parameter k untuk menampung berapa fitur yang ingin diambil dan untuk penilitian ini menggunakan k = 40.

c. Scalling Data

Pada tahapan scaling data, maka semua fitur akan memiliki skala yang sama sehingga algoritma akan mengambil keputusan dengan lebih objektif. Dalam penelitian ini untuk melakukan scaling data, menggunakan StandardScaler().

Berikut ini contoh 1 data pertama yg telah di scaling dengan menggunakan StadardScaler().

```
array([[ 3.11853094, -0.91073638, -0.62455876, -0.89034938, -0.60010573,
        -0.74924866, -0.777536  , -0.74157241,  1.31675643,  1.01731211,
        -1.01479735, -1.02096249, -0.96575752, -0.26220256, -0.8399904 ,
        -0.82087172, -0.80249712, -0.82558735, -0.22633361, -0.67281875,
        -0.66219368, -0.68157476, -0.66499342, -0.17060101,  1.01633668,
        -1.38169285, -0.75350021, -0.79616009, -0.7901402 , -0.78478765,
        -0.80776941, -0.89034938, -0.7775382 , -0.6245581 , -0.30306894,
        -0.1988214 , -0.29837001, -0.65467354, -0.66631824, -0.65573597],
```

Gambar 3. Data yang telah dilakukan StandardScaler()

Pada tahapan ini maka data telah bersih setelah di pre-processing dan hasilnya siap untuk diproses pada tahap selanjutnya yaitu tahap *modeling*.

3. *Modeling*

Pada tahap *modeling* akan dilakukan beberapa langkah untuk sampai pada tahap evaluasi. Berikut langkah langkah yang dilakukan :

a. *Train test* split, pembagian data *train* dan *test* dengan proporsi *test* 20% dan 80% sebagai data *train*, 10% untuk data *test* dan 80% data *train* serta 30% data *test* dan 70% data *train* sebagai pembanding.

b. Modelling dan evaluation

Saat tahap pemodelan dilakukan, digunakan algoritma *naïve bayes* dengan perbandingan model GaussianNB dan BernoulliNB yang menghasilkan akurasi sebagai berikut :

Model	Train_test_split	Accuracy
GaussianNB	9:1	83.41%
BernoulliNB	9:1	76.28%
GaussianNB	8:2	83.42%
BernoulliNB	8:2	76.20%
GaussianNB	7:3	83.45%
BernoulliNB	7:3	76.21%

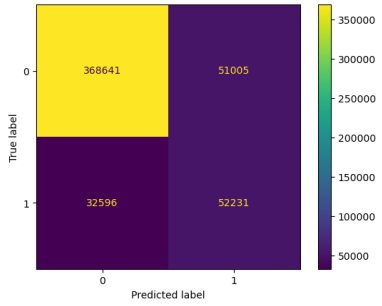
Gambar 6. Perbandingan Model

Dapat dilihat pada gambar diatas bahwa algoritma GaussianNB menghasilkan akurasi yang paling tinggi dengan parameter yang ditunjukkan pada gambar model dan parameter dibawah ini :

```
GaussianNB()
{'priors': None, 'var_smoothing': 1e-09}
```

Gambar 7. Model dan Parameter

Kemudian, pada tahap evaluasi model, ditemukan bahwa hasil akurasi untuk metode *Naïve Bayes* adalah 83,45% dengan menggunakan data *train* 70% serta data *test* 30%. Hasil dari implementasi metode *Naïve Bayes* menggunakan algoritma GaussianNB untuk deteksi Anomali dan *Low Rate DDoS* dapat dilihat pada confusion matriks yang di tampilkan pada gambar dibawah ini :



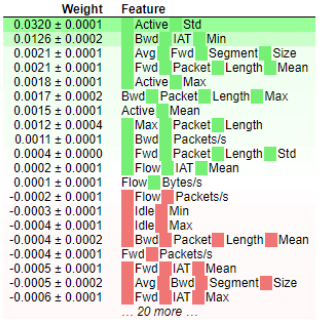
Gambar 8. *Confusion Matrix Naïve Bayes*

```
acccuracy score : 83.42805264107295 %
precision score : 61.57355558961179 %
Recall score : 50.593785113720024 %
f1 score : 55.54627970414171 %
```

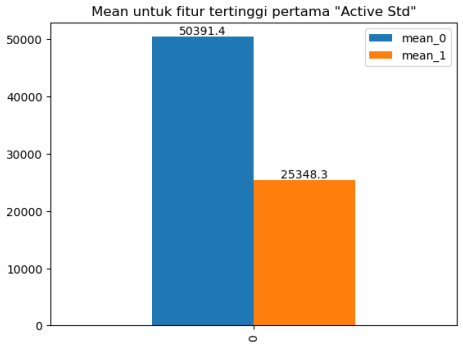
Gambar 9. Hasil Akurasi

4. Data *storytelling*

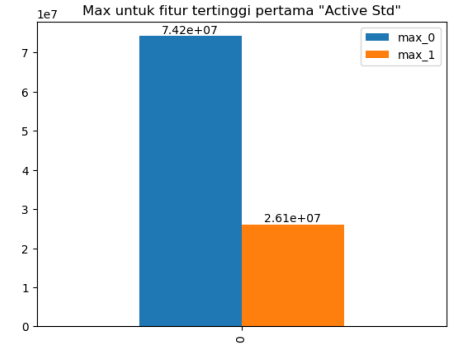
Pada tahap ini akan menjelaskan tentang 3 fitur terbaik dari model yang telah dibuat berdasarkan *feature importance* yang dihasilkan. Hal tersebut dapat dilihat pada gambar dibawah ini :



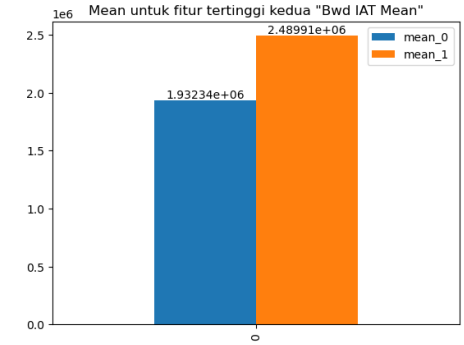
Gambar 10. *Feature Importance*



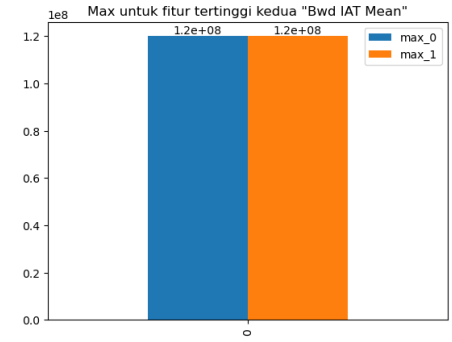
Gambar 11. Mean untuk Fitur Tertinggi Pertama



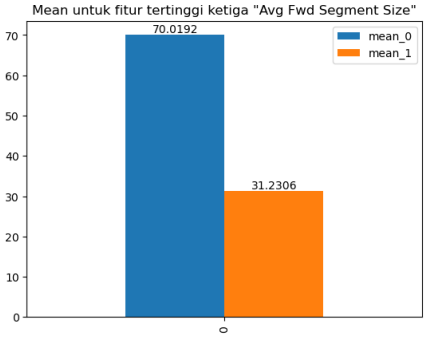
Gambar 12. Max untuk Fitur Tertinggi Pertama



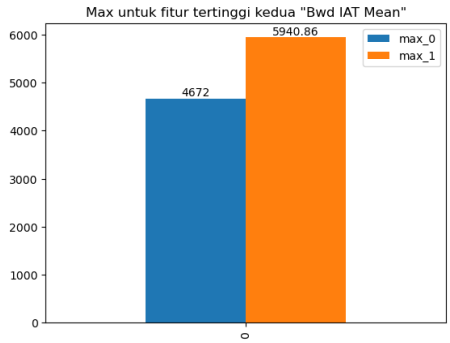
Gambar 13. Mean untuk Fitur Tertinggi Kedua



Gambar 14. Max untuk Fitur Tertinggi Kedua



Gambar 15. Mean untuk Fitur Tertinggi Ketiga



Gambar 16. Max untuk Fitur Tertinggi Ketiga

V. KESIMPULAN

Berdasarkan penelitian yang dilakukan dalam mengeksplorasi dan mengevaluasi model tentang mendeteksi anomali dan *Low Rate DDoS* dalam jaringan lalu lintas, hasil dari penelitian ini menunjukkan bahwa pemilihan fitur untuk mendeteksi anomaly dan *Low Rate DDoS* dapat menggunakan chisquare atau di dalam python adalah chi2. Sehingga fitur menjadi lebih spesifik dan bisa lebih dipahami oleh model *Machine Learning*. Diharapkan untuk penelitian selanjutnya bisa menggunakan metode yang lain dalam melakukan *feature selection* agar lebih spesifik fitur yang dihasilkan.Hasil akurasi yang paling tinggi terdapat pada model GaussianNB yaitu dengan akurasi 83,45 % dimana telah dibandingkan oleh model BernoulliNB yang tertinggi hanya mendapatkan akurasi 76,21%. Adapun Precision dan F1 Score pada model GaussianNB adalah 61.573 % dan 55.546 %. Dengan seperti ini, maka model *Machine Learning* GaussianNB bisa dikatakan sudah cukup baik dalam mengenali pola dan mendeteksi anomali dan *Low Rate DDoS* dalam jaringan lalu lintas. Diharapkan penelitian selanjutnya bisa mengembangkan model dengan algoritma ensemble-method agar model lebih meningkat akurasi yang dihasilkan.

REFERENSI

[1] Eskol F, Sirait T. Dampak Revolusi Industri 4.0 pada Industri Teknologi Komunikasi di Indonesia: Peluang dan Tantangan. Jurnal Penelitian dan Pengembangan Sains dan Humaniora [Internet]. 2022;6(1):132–9. Available from: https://doi.org/10.23887/jppsh.v6i1.28153

[2] Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya Jl Raya Palembang-Prabumulih Km J, Ogan Ilir I, Ashari A, Setiawan H, Ilmu Komputer dan Elektronika J, Mipa F, et al. Cloud Computing : Solusi ICT ? Jurnal Sistem Informasi (JSI) [Internet]. 2011;3(2):336–45. Available from: http://ejournal.unsri.ac.id/index.php/jsi/index

[3] Lin H, Wu C, Masdari M. A comprehensive survey of network traffic anomalies and *DDoS* attacks detection schemes using fuzzy techniques. Computers and Electrical Engineering. 2022 Dec 1;104:108466.

[4] Ahalawat A, Babu KS, Turuk AK, Patel S. A low-rate *DDoS* detection and mitigation for SDN using Renyi Entropy with Packet Drop. Journal of Information Security and Applications. 2022 Aug 1;68.
[5] Miladinović A, Iscra K, Ajčević M, Restivo L, Kresevic S, Merlo M, et al. Naive Bayesian-based nomogram for identification of early asymptomatic Dilated Cardiomyopathy.
[6] Mayang Sari I, Rahman Wijaya D, Hidayat W. BERBASIS *DATASET* ELECTRONIC NOSE MENGGUNAKAN ALGORITMA *NAÏVE BAYES* CLASSIFIER. Proceeding of Applied Science. 2021;7(6):2589.
[7] Sibarani NS, Munawar G, Wisnuadhi B. Analisis Performa Aplikasi Android Pada Bahasa Pemrograman Java dan Kotlin.
[8] Kurniabudi, Stiawan D, Darmawijoyo, Bin Idris MY Bin, Bamhdi AM, Budiarto R. CICIDS-2017 *Dataset* Feature Analysis with Information Gain for Anomaly Detection. IEEE Access. 2020;8:132911–21.