Jurnal Pendidikan dan Konseling

Volume 4 Nomor 6 Tahun 2022

E-ISSN: 2685-936X dan P-ISSN: 2685-9351

Universitas Pahlawan Tuanku Tambusai

Optimasi Metode Naïve Bayes dengan *Particle Swarm Optimization* untuk Sistem Deteksi Serangan D-Dos

Doni Syahroni[1], Nurjaya[2]

Mahasiswa Program Studi Teknik Informatika Fakultas Teknik Universitas Pamulang[1],

Program Studi Teknik Informatika Fakultas Teknik Universitas Pamulang[2]

Email : donihamster88@gmail.com[1], dosen00370@unpam.ac.id [2]

Abstrak

D-DoS (Distributed Denial of Service) adalah jenis serangan terstruktur. Serangan DDoS mampu melumpuhkan server dengan membanjiri lalu lintas jaringan dan mengakibatkan down[11]. Ancaman dan serangan terhadap keamanan server terus meningkat. Kepala Badan Siber dan Sandi Negara (BSSN) Hinsa Siburian mengatakan selama 2021 tercatat ada 888.711.736 serangan siber[5]. Telah banyak dilakukan penelitian tentang pendeteksian serangan DDoS dengan berbagai metode, seperti Artificial Neural Network (ANN)[12], K-Nearest neighbor dengan optimasi menggunakan Principal Component Analysis (PCA)[15], dan Naïve Bayes[8]. Namun metode Nave Bayes memiliki kelemahan yaitu tidak dapat melakukan seleksi atribut dan belum ada yang mengoptimasi klasifikasi Naïve Bayes menggunakan Particle Swarm Optimization dalam memberikan bobot pada setiap atribut untuk mendeteksi serangan D-DoS. Klasifikasi Naïve Bayes dengan melakukan optimasi menggunakan Particle Swarm Optimization untuk mengetahui seberapa besar nilai akurasi sebelum dan sesudah dilakukan optimasi. Berdasarkan penelitian yang telah dilakukan dengan menggunakan metode Naïve Bayes Classification (NBC) yang belum dioptimasi menunjukkan akurasi sebesar 0.9178 atau 92%, MSE 0.0821 atau 0.08% dan AUC 0.9221 atau 92% dan telah di optimasi Particle Swarm Optimization (PSO) dengan hasil menunjukkan akurasi sebesar 0,948 atau 95%, MSE 0,0516 atau 0,05% dan AUC 0,9463 atau 95%. Oleh karena itu, dengan optimalisasi menggunakan metode PSO, proses pendeteksian serangan D-DoS mengalami peningkatan sebesar 3%

Kata Kunci : *Mengklasifikasi Naïve Bayes; Optimasi Kawanan Partikel; Klasifikasi; Pembobotan Atribut; Deteksi D-DoS.*

Abstract

D-DoS (Distribute Denial of Service) is a type of structured attack. DDoS attacks are able to paralyze servers by flooding network traffic and resulting in downs[11]. Threats and attacks on server security are constantly increasing. Head of the National Cyber and Crypto Agency (BSSN) Hinsa Siburian said that during 2021 there were 888,711,736 cyber attacks recorded[5]. Many have conducted research on the detection of DDoS attacks with various methods, such as Artificial Neural Network (ANN)[12], K-Nearest neighborhood with optimization using Principal Component Analysis (PCA)[15], and Naïve Bayes[8]. However, the Nave Bayes method has a weakness, namely it cannot perform attribute selection and no one has optimized the Naïve Bayes classification using Particle Swarm Optimization in assigning weights to each attribute to detect D-DoS attacks. Naïve Bayes classification by optimizing using Particle Swarm Optimization to find out how much the accuracy value is before and after optimization. Based on research that has been carried out using the Naïve Bayes Classification (NBC) method that has not been optimized, it shows an accuracy of 0.9178 or 92%, MSE 0.0821 or 0.08% and

AUC 0.9221 or 92% and has been optimized by Particle Swarm Optimization (PSO) with results showing accuracy of 0.948 or 95%, MSE 0.0516 or 0.05% and AUC 0.9463 or 95%. Therefore, by optimizing using the PSO method, the D-DoS attack detection process has increased by 3%
Keywords *: Naïve Bayes Classifier; Particle Swarm Optimization; Classification; Attribute Weighting; D-DoS detection.*

PENDAHULUAN

DDoS (Distribute Denial of Service) merupakan jenis serangan yang terstruktur. Serangan DDoS mampu melumpuhkan server dengan membanjiri lalu lintas jaringan dan mengakibatkan down. Ancaman dan serangan terhadap keamanan server terus meningkat. Kepala Badan Siber dan Sandi Negara (BSSN) Hinsa Siburian mengatakan, selama tahun 2021 ini tercatat ada 888.711.736 serangan siber. Adapun data tersebut adalah merupakan data yang dihimpun sejak Januari hingga Agustus 2021 lalu. Oleh karena itu, Presiden Joko Widodo mengingatkan agar Indonesia selalu bersiap menghadapi serangan siber dan kejahatan penggunaan data. Kemudian ada juga yang berpendapat bahwa DDoS merupakan salah satu jenis serangan Denial of Service dimana serangan ini menggunakan banyak host atau bisa disebut dengan komputer zombie yang bertujuan untuk menyerang secara bersamaan dengan mengirimkan data atau request secara berulang-ulang dengan tujuan agar komputer target tidak dapat berfungsi dengan baik dan jaringannya akan terganggu.

Beberapa metode untuk mendeteksi serangan D-DoS yang telah dilakukan oleh para peneliti yaitu dengan menggunakan algoritma Artificial Neural Network (ANN), K-Nearest Neighborhood dengan optimasi menggunakan Principal Component Analysis (PCA) dan Naïve Bayes.

Beberapa metode penelitian mengenai sistem deteksi serangan DDoS telah dilakukan, penelitian yang dilakukan oleh Ahmad Sanmorino yang berjudul A study for DDOS attack classification method di tahun 2019. Dalam penelitian ini menggunakan tiga algoritma yaitu Decisition Tree, Naïve Bayes , dan Artificial Neural Network yang bertujuan untuk mengetahui algoritma yang terbaik dalam mendeteksi serangan D-DoS. Hasil dalam penelitian yang dilakukan oleh Ahmad 2 Sanmorino bahwa algoritma Artificial Neural Network (ANN) lebih tinggi akurasinya sebesar 84.5% (True Positive) dari pada algoritma Naïve Bayes sebesar 76.6% (True Positive) dan Decision Tree sebesar 84.0% (True Positive) karena ANN mampu mengetahui pola serangan D-DoS yang tidak dapat diprediksi atau selalu berubah dan tidak memiliki parameter yang jelas. Tetapi kelemahan dari ANN ini yaitu dalam memproses operasi dan training jika jumlah data yang besar dan tidak adanya aturan khusus dalam menentukan struktur ANN, sehingga menghasilkan akurasi klasifikasi yang belum maksimal.

Selanjutnya penelitian yang dilakukan oleh S. Umarani dan D. Sharmila yang berjudul Predicting Application Layer DDoS Attacks Using Machine Learning Algorithms pada tahun 2014. Dalam penelitian ini menggunakan dataset yang berisikan access logs in the 1998 World Cup Web site dari 30 April 1998 – 26 Juli 1998. Algoritma yang digunakan pada penelitian ini yaitu Naïve Bayes dan K-Nearest Neighborhood yang dioptimasi menggunakan algoritma Principal Component Analysis (PCA). Hasil penelitian ini mampu mendapatkan nilai detection rate Naïve Bayes 95% dan KNN 93%, ketika sudah dioptimasi oleh PCA menjadi Naïve Bayes 96% dan KNN 94%, oleh karena itu detection rate tertinggi yaitu NBC+PCA. [7]Karena kelebihan dari algoritma PCA ini yaitu mampu mengurangi jumlah atribut yang kurang baik , tetapi algoritma PCA ini kurang optimal dalam pemisahan antar kelas.

Kemudian penelitian yang dilakukan oleh Arief Prasetyo, Luqman Affandi dan Dedi Arpandi yang berjudul Implementasi Metode Naïve Bayes Untuk Intrusion Detection System (IDS) di tahun 2018. Dalam penelitian ini dataset yang digunakan adalah data NSL-KDD, NSL-KDD telah menyediakan data training dan data testing untuk proses penelitian klasifikasi serangan. Dari dataset tersebut akan dilakukan klasifikasi serangan menggunakan naïve bayes. Tetapi field dari dataset NSL-KDD tidak

semuanya digunakan, hanya 8 field saja yaitu 8 field tersebut yaitu, src_bytes, dst_bytes, count, srv_count, dst_host_count, dst_host_srv_count, dst_host_same_src_port_rate, dan dst_host_srv_diff_host_rate. Jumlah data testing yang dilakukan yaitu 100 – 200 data dan jumlah data training yaitu 1500 – 5000 data. Penelitian ini berhasil melakukan klasifikasi serangan-serangan baru 3 dengan akurasi kebenaran adalah sebesar 81- 84,67 % . Algoritma ini memiliki keakuratan yang cukup baik diterapkan pada data yang besar dan dapat menangani data yang tidak lengkap (missing value) serta kuat terhadap atribut yang tidak relevan dan noise pada data, tetapi NBC tidak dapat melakukan seleksi atribut sehingga dapat mempengaruhi nilai akurasinya.

Berdasarkan penelitian di atas untuk memprediksi serangan D-DoS masing masing memiliki kekurangan di setiap algoritma yang digunakan. Namun, peneliti menemukan masalah yang dapat diselesaikan yaitu pada penelitian yang dilakukan oleh peneliti-peneliti sebelumnya. Pada penelitian tersebut terdapat kurang optimalnya akurasi yang dihasilkan oleh algoritma Naïve Bayes, karena tidak dapat melakukan seleksi atribut. Oleh karena itu, belum ada yang menggunakan Particle Swarm Optimization untuk melakukan optimasi pada klasifikasi Naive Bayes dengan cara pemberian bobot pada setiap atribut (attribute weight). Dalam penelitian yang akan peneliti lakukan yaitu melakukan optimasi pada klasifikasi Naïve Bayes menggunakan Particle Swarm Optimization untuk sistem deteksi serangan D-DoS.

Adapun tujuan dari penelitian ini adalah meningkatkan akurasi klasifikasi naïve bayes dengan melakukan optimasi menggunakan Particle Swarm Optimization untuk mengetahui berapa besarnya nilai akurasi sebelum di optimasi dan setelah di optimasi.

Data *Mining*

Data mining adalah sebuah proses untuk mendapatkan informasi yang berharga dari sekumpulan data yang sangat besar dengan menganalisa data menggunakan algoritma atau metode yang dapat dipakai untuk kepentingan organisasi, pribadi ataupun perusahaan.

Naïve Bayes *Classifier*

Naive Bayes Classifier (NBC) merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Metode NBC menempuh dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan dilakukan proses analisis terhadap sampel dokumen berupa pemilihan vocabulary atau kumpulan kosakata, yaitu kata yang mungkin muncul dalam koleksi dokumen sampel yang sedapat mungkin dapat menjadi representasi dokumen. Selanjutnya adalah penentuan probabilitas prior bagi tiap kategori berdasarkan sampel dokumen. Pada tahap klasifikasi ditentukan nilai kategori dari suatu dokumen berdasarkan term yang muncul dalam dokumen yang diklasifikasi.

Umumnya, Bayes mudah dihitung untuk fitur bertipe kategoris, nama untuk fitur dengan tipe numerik (kontinu) ada perlakukan khusus sebelum dimasukkan dalam Naïve Bayes , adalah sebagai berikut:

a. Melakukan diskritisasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskrit. Pendekatan ini dilakukan dengan mentransformasi fitur kontinu ke dalam fitur ordinal.
b. Mengasumsikan bentuk tertentu dari distribusi probabilitas untuk fitur continue dan memperkirakan parameter distribusi dengan data pelatihan. Distribusi Gaussian biasanya dipilih untuk merepresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas P(Xi|Yi), sedangkan distribusi Gaussian dikarakteristikkan dengan dua parameter: mean (μ) dan varian (σ 2). Untuk setiap kelas Yj , probabilitas bersyarat kelas Yj untuk fitur Xi adalah :

$$P(X_i = X_i | Y = Y_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{-\frac{(x_j - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (2.1)$$

Keterangan :

μ_{ij} = Didapat dari mean sampel Xi ($\overline{x}$) / rata rata semua data latih yang menjadi milik kelas Yj

σ_{ij}^2 = Varian sampel (s^2) dari data latih yang menjadi kelas Yj

e = eksponensial (2.71828183)

Untuk rumus mean / rata – rata sebagai berikut :

$$\underline{x} = \frac{Xi_1 | Y_j + Xi_2 | Y_j + Xi_3 | Y_j + \cdots Xi_n | Y_j}{n} \quad (2.2)$$

Keterangan :

Xi | Yj = Nilai Xi terhadap kelas Yj

n = Banyaknya Xi ke Yj

Untuk rumus varian pangkat 2 (σ^2) / sebagai berikut :

$$s^2 | Yj = \frac{(Xi_1 | Y_j - \underline{x})^2 + (Xi_2 | Y_j - \underline{x})^2 + \cdots (Xi_n | Y_j - \underline{x})^2}{n-1} \quad (2.3)$$

Keterangan :

Xi | Yj = Nilai Xi terhadap kelas Yj

$\underline{x}$ = Mean / rata rata atribut

n = Banyaknya Xi ke Yj

Untuk rumus varian (σ) / sebagai berikut :

$\sqrt{\sigma^2}$ (2.4)

Keterangan :

σ^2 = Hasil dari perhitungan pada rumus varian pangkat 2 atau $s^2 | Yj$

Particle Swarm Optimization

Particle Swarm Optimization (PSO) adalah salah satu dari teknik komputasi evolusioner, yang mana populasi pada PSO didasarkan pada penelusuran algoritma dan diawali dengan suatu populasi yang random yang disebut dengan particle. Berbeda dengan teknik komputasi evolusioner lainnya, setiap particle di dalam PSO juga berhubungan dengan suatu velocity. Particle-particle tersebut bergerak melalui penelusuran ruang dengan velocity yang dinamis yang disesuaikan menurut perilaku historisnya. Oleh karena itu, particle-particle mempunyai kecenderungan untuk bergerak ke area penelusuran yang lebih baik setelah melewati proses penelusuran.

Aliran algoritma PSO dimulai dengan jumlah populasi partikel yang posisinya merupakan solusi potensial untuk masalah yang tengah dipelajari dan kecepatan yang secara acak diinisialisasi dalam ruang pencarian. Pada setiap iterasi, pencarian untuk posisi yang optimal dilakukan dengan memperbarui kecepatan dan posisi partikel. Juga pada setiap iterasi, nilai fitness setiap posisi partikel ditentukan menggunakan fungsi fitness. Kecepatan masing-masing partikel diperbarui menggunakan dua posisi terbaik, posisi personal terbaik dan posisi global terbaik. Posisi personal terbaik, pbest adalah posisi terbaik partikel yang telah dikunjungi dan gbest adalah posisi terbaik yang telah dikunjungi swarm sejak langkah pertama kalinya. Jika solusi local best mempunyai suatu biaya yang kurang dari biaya solusi global yang ada, maka solusi local best menggantikan solusi global best.

Berikut ini merupakan formulasi matematika yang menggambarkan posisi dan kecepatan partikel pada suatu dimensi ruang tertentu :

$X_i(t) = x_{i1}(t), x_{i2}(t), \dots, x_{iN}(t)$ (2.5)

$V_i(t) = v_{i1}(t), v_{i2}(t), \dots, v_{iN}(t)$ (2.6)

Dimana :

X = posisi partikel

V = kecepatan partikel

i = indeks partikel

t = iterasi ke-t

N = ukuran dimensi ruang

Berikut ini merupakan model matematika yang menggambarkan mekanisme updating status partikel Kennedy and Eberhart [1995]:

$V_i(t) = V_i(t-1) + c_1 r_1 \left(X_i^L - X_i(t-1)\right) + c_2 r_2 \left(X^G - X_i(t-1)\right)$ (2.7)

$X_i(t) = V_i(t) + X_i(t-1)$ (2.8)

Dimana :

$X_i^L = X_{i_1}^L, X_{2,}^L \dots, X_{i_N}^L$ merepresentasikan local best dari partikel ke-i. Sedangkan

$X^G = X_{i_1}^G, X_{2,}^G \dots, X_{i_N}^G$ merepresentasikan global best dari seluruh kawanan.

X = Posisi Partikel

V = Merupakan kecepatan (velocity)

t = iterasi ke-t

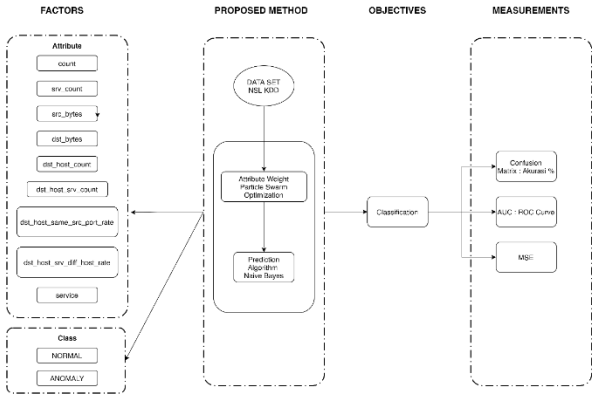
i = index partikel

Sedangkan c_1 dan c_2 adalah suatu konstanta yang bernilai positif yang biasanya disebut sebagai learning factor. Kemudian r_1 dan r_2 adalah suatu bilangan random yang bernilai antara 0 sampai 1. Persamaan 2.7 digunakan untuk menghitung kecepatan partikel yang baru berdasarkan kecepatan sebelumnya, jarak antara posisi saat ini dengan posisi terbaik partikel (local best), dan jarak antara posisi saat ini dengan posisi terbaik kawanan (global best). Kemudian partikel terbang menuju posisi yang baru berdasarkan persamaan 2.8. Setelah algoritma PSO ini dijalankan dengan sejumlah iterasi tertentu hingga mencapai kriteria pemberhentian, maka akan didapatkan solusi yang terletak pada global best.

D-DoS Attack

Pengertian dari DDoS yaitu serangan yang dilakukan oleh hacker untuk mengganggu pengguna jaringan dengan cara membanjiri lalu lintas data yang tinggi terhadap server, sehingga server tidak dapat lagi memberikan layanan dan terjadi hang pada server.

Kerangka Pemikiran



Gambar 1 Kerangka Pemikiran

Dalam kerangka pemikiran diatas dijelaskan bahwa penelitian ini ditujukan untuk mengetahui seberapa meningkatkan akurasi naïve bayes dengan menggunakan particle swarm optimization dalam memprediksi serangan DDoS dan menghasilkan nilai akurasi , AUC dan MSE.

METODE

Rancangan Penelitian

Dalam penelitian ini metode yang digunakan dalam menerapkan algoritma *Particle Swarm Optimization* untuk meningkatkan hasil prediksi algoritma Naïve Bayes dalam mendeteksi serangan D-DoS adalah metode eksperimental, dimana data yang diambil merupakan dataset NSL-KDD dari *UNB (University Of New Brunswick)* hasil penelitian yang sudah dilakukan sebelumnya atau data sekunder.

Pengumpulan Data

Data yang digunakan untuk melakukan penelitian adalah dataset NSL-KDD yang berasal dari *UNB (University Of New Brunswick)*.

Pengolahan Data Awal

Dalam data set NSL-KDD ini memiliki empat kategori serangan yang terjadi[8], yaitu DoS (*Denial of Service), PROBING,* R2L (*Remote to Local),* dan U2R (*User to Root).*

Pada pengolahan data awal ini dilakukan pemilihan atribut atau tahap preprocessing yang akan digunakan dalam penelitian ini yang sebelumnya sebanyak 25192 data dengan 41 atribut dan 1 kelas menjadi sebanyak 22495 dengan 9 atribut dan 1 kelas yang diperlukan dalam pendeteksian serangan D-DoS yaitu *service, src_bytes, dst_bytes, count, srv_count, dst_host_count, dst_host_srv_count, dst_host_same_src_port_rate, dst_host_srv_diff_host_rate* dan *class*.

Dikarenakan dalam penelitian ini peneliti hanya membahas serangan D-DoS, maka data yang peneliti ambil hanya kategori DoS saja.

Setelah data tersebut di filter sesuai dengan kategori serangan D-DoS , maka selanjutnya mengubah nilai pada atribut *service* dari kata menjadi angka, agar data dapat diproses menggunakan teknik *preprocessing data label encoding.* Dan untuk class atau nilai y juga diubah menjadi angka, namun hanya dibagi menjadi 2 *output* yaitu 0 dan 1. Nilai 0 merupakan "*normal*" dan nilai 1 selain "*normal*",

Setelah proses *label encoding* tersebut dilakukan, maka selanjutnya peneliti akan menghapus data yang duplikat untuk mengetahui pola data yang digunakan. Berikut hasil proses sebelum dan sesudah penghapusan pola data yang duplikat :

Tabel 1 Total dataset sebelum dan sesudah proses penghapusan data duplikat

No	Sebelum (total data)	Sesudah (total data)
1	22495	21782

Dan dibawah ini adalah sampel dari pola data yang akan peneliti gunakan:

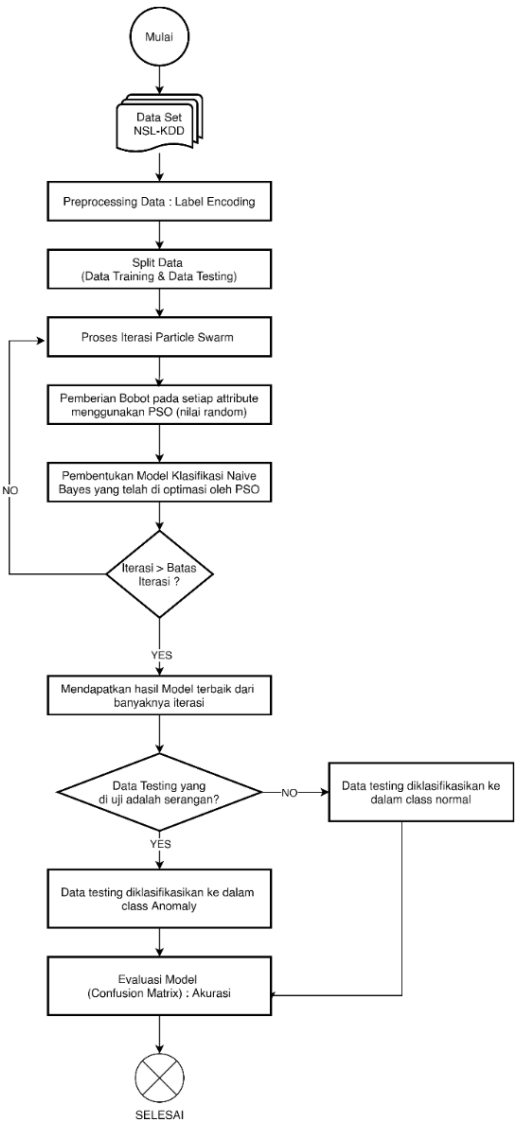
Tabel 2 Pola data yang digunakan

No	x0	x1	x2	x3	x4	x5	x6	x7	x8	T
1	0	0	0	230	14	255	14	0	0	1
2	0	9	38	2	1	255	13	0	0	0
3	0	0	0	266	9	255	9	0	0	1
4	0	10	35	1	1	255	10	0	0	0
…										
21782	63	0	0	271	8	255	8	0	0	1

Metode Yang Diusulkan

Dalam penelitian ini metode yang digunakan adalah *Particle Swarm Optimization* yang bertujuan untuk melakukan optimasi hasil prediksi terhadap algoritma *Naïve Bayes* dalam memprediksi serangan D-DoS.

Untuk melihat lebih jelas tentang metode yang diusulkan dalam penelitian ini dapat dilihat pada gambar dibawah ini:



Gambar 1 Metode yang diusulkan

Langkah Menyelesaikan Metode

Langkah penyelesaian metode dalam penelitian ini dilakukan dengan beberapa tahapan, sebagai berikut :

a. Masukkan *dataset NSL-KDD dan* dilakukan pemisahan data untuk mendapatkan data latih dan data uji.
b. Penentuan nilai parameter besar populasi dan jumlah maksimum generasi yang tepat dengan memperhatikan waktu komputasi.
c. Melakukan inisialisasi partikel dengan posisi acak (nilai random) dan vektor kecepatan (*velocity*) diatur nilai awal adalah 0 pada setiap attribute. *Local Best* awal diambil dari posisi partikel dengan *fitness* terbaik, dan *Global Best* awal diset sama dengan *Local Best* (Peneliti menggunakan *library PySwarm*);
d. Hitung dan update kecepatan partikel pada setiap iterasi , dengan rumus :

$$V_i(t) = V_i\,(t-1) + c_1 r_1 \left(X_i^L - X_i\,(t-1)\right) + c_2 r_2 \left(X^G - X_i\,(t-1)\right) \quad (3.1)$$

$$X_i(t) = V_i(t) + X_i\,(t-1) \quad (3.2)$$

e. Evaluasi nilai fitness untuk setiap posisi partikel (P). Jika fitness (P) lebih baik dari pada fitness O

(Pbest), maka Local Best (Pbest) = P;

f. Tentukan Pbest terbaik sebagai Gbest dan rumus prediksi yang digunakan sebagai pembentukan model, Jika posisi partikel tersebut lebih baik daripada fitness terbaik sebelumnya, maka posisi tersebut dijadikan sebagai Global Best (Gbest);

g. Jika Gbest adalah solusi optimal maka nilai tersebut adalah bobot pada setiap nilai atribut dan akan dijadikan model data training atau data latih dengan nilai akurasi terbaik. Jika belum optimal maka dilakukan iterasi hingga batas maksimum iterasi;

h. Dalam pembuatan model data latih rumus yang akan digunakan berbeda dengan rumus naïve bayes klasifikasi sebelumnya, karena ada proses perhitungan bobot yang akan dikalikan dengan setiap attribute dalam mencari mean dan varian.

Pada rumus mean / rata – rata sebagai berikut :

$$\underline{x} = \frac{Xi_1 \mid Y_j + Xi_2 \mid Y_j + Xi_3 \mid Y_j + \cdots Xi_n \mid Y_j}{n} \; w_i \quad (3.3)$$

Keterangan :

Xi | Yj = Nilai Xi terhadap kelas Yj

n = Banyaknya Xi ke Yj

w_i = Bobot yang didapatkan dari *Particle Swarm Optimization* ke i

Pada rumus varian pangkat 2 (σ^2) / sebagai berikut :

$$s^2 \mid Yj = \frac{(Xi_1 \mid Y_j - \underline{x})^2 + (Xi_2 \mid Y_j - \underline{x})^2 + \cdots (Xi_n \mid Y_j - \underline{x})^2}{n-1} \; w_j \quad (3.4)$$

Keterangan :

Xi | Yj = Nilai Xi terhadap kelas Yj

$\underline{x}$ = Mean / rata rata atribut

n = Banyaknya Xi ke Yj

w_i = Bobot yang didapatkan dari *Particle Swarm Optimization* ke i

Dan terakhir pembentukan klasifikasi menggunakan algoritma Naïve Bayes yang sudah di optimasi dengan Particle Swarm Optimization untuk menentukan class normal atau anomaly.

HASIL DAN PEMBAHASAN

Perhitungan Pada Metode NBC+PSO

Pada perhitungan Naïve Bayes Klasifikasi yang telah dioptimasi oleh *Particle Swarm Optimization* hampir sama dengan langkah-langkah perhitungan Naïve Bayes klasifikasi pada umumnya, yang membedakan adalah saat perhitungan mendapatkan hasil *mean* dan *varian* akan dikalikan dengan bobot yang didapatkan dari PSO, contohnya pada X0 dengan bobot 2.67213798 sebagai berikut

Hitung *mean :*

$$\underline{X0}_{yes} = \frac{28+42+\cdots+3}{6870} = 180.7561 \quad (4.1)$$

= 180.7561 x 2.67213798 = 483.005244

$$\underline{X0}_{no} = \frac{20+9+\cdots}{10555} = 22.6971 \quad (4.2)$$

= 22.6971 x 2.67213798 = 60.6497829

Hitung *varian :*

$$S0^2{}_{yes} = \frac{(28-180.7561)^2+(42-180.7561)^2+\cdots+(3-180.7561)^2}{6870-1} = 10118.9163 \quad (4.3)$$

= 10118.9163 x 2.67213798 = 27039.1405

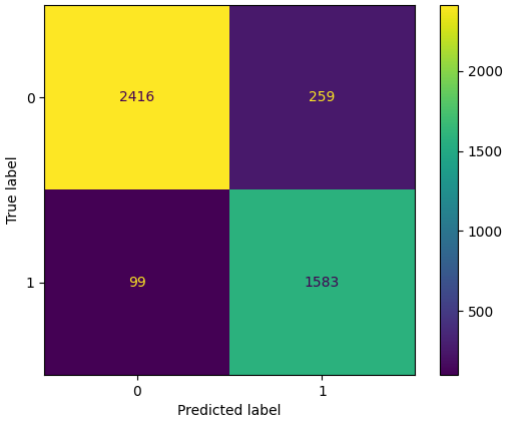
$$S0^2{}_{no} = \frac{(20-22.6971)^2+(9-22.6971)^2+\cdots}{10555-1} = 2861.8141 \quad (4.4)$$

= 2861.8141 x 2.67213798 = 7647.1621

$$S0_{yes} = \sqrt{27039.1405} = 164.4358 \quad (4.5)$$
$$S0_{no} = \sqrt{7647.1621} = 87.4480 \quad (4.6)$$

1. Evaluasi Hasil

Hasil eksperimen dan pengujian model menggunakan NBC :



Gambar 2 *Confusion Matrix* Hasil NBC

Gambar di atas menerangkan bahwa data prediksi dan data kebenarannya (aktual) menghasilkan nilai, sebagai berikut :

a. *True Positive* dengan nilai 1563, artinya hasil prediksi bernilai 1 dan data aktual bernilai 1;
b. *True Negative* dengan nilai 2416, artinya hasil prediksi bernilai 0 dan data aktual bernilai 0;
c. *False Positive* dengan nilai 259, artinya hasil prediksi bernilai 1 tetapi data aktual bernilai 0;
d. *False Negative* dengan nilai 99, artinya hasil prediksi bernilai 0 tetapi data aktual bernilai 1;

Dari hasil tersebut, mendapatkan nilai akurasi sebesar 0.9178 , MSE (Mean Squared Error) sebesar 0.0821, Precision sebesar 0.8593, Recall sebesar 0.9411 dan AUC (Area Under Curve) sebesar 0.9221.

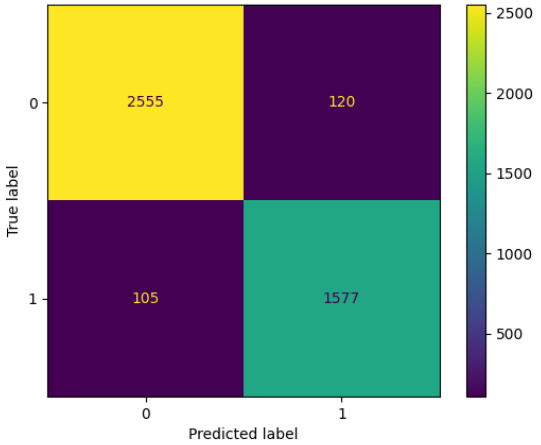
Hasil eksperimen dan pengujian model menggunakan NBC + PSO :

Tabel 3 Hasil Evaluasi PSO

Iterasi	Jumlah Partikel	Nilai Akurasi
10	5	0.9341
15	5	0.9483
20	5	0.9008
25	5	0.8941
30	5	0.8909
40	5	0.8909
50	5	0.8957

Tabel diatas menerangkan bahwa dari partikel di setiap iterasi mendapatkan berbagai nilai

akurasi dalam memprediksi serangan D-DoS, dan akurasi terbaik ada pada iterasi ke 15 sebesar 0.9483 atau 95%. Jika, dilihat dalam confusion matrix hasilnya sebagai berikut :



Gambar 3 *Confusion Matrix* Hasil NBC+PSO

Gambar di atas menerangkan bahwa data prediksi dan data kebenarannya (aktual) menghasilkan nilai, sebagai berikut :

a. *True Positive* dengan nilai 1577, artinya hasil prediksi bernilai 1 dan data aktual bernilai 1;
b. *True Negative* dengan nilai 2555, artinya hasil prediksi bernilai 0 dan data aktual bernilai 0;
c. *False Positive* dengan nilai 120, artinya hasil prediksi bernilai 1 tetapi data aktual bernilai 0;
d. *False Negative* dengan nilai 105, artinya hasil prediksi bernilai 0 tetapi data aktual bernilai 1;

Dari hasil tersebut, mendapatkan nilai akurasi sebesar 0.9483 , MSE (Mean Squared Error) sebesar 0.0516, Precision sebesar 0.9292, Recall sebesar 0.9375 dan AUC (Area Under Curve) sebesar 0.9463.

SIMPULAN

Dalam penelitian ini dilakukan optimasi klasifikasi Naïve Bayes menggunakan Particle Swarm Optimization untuk mendeteksi serangan D-DoS.

Berdasarkan penelitian yang telah dilakukan dalam memprediksi serangan D-DoS pada data dari dataset NSL-KDD dengan metode Naïve Bayes Classification (NBC) yang belum teroptimasi menunjukan hasil akurasi sebesar 0.9178 atau 92%, MSE 0.0821 atau 0.08% dan AUC 0.9221 atau 92% dan telah teroptimasi oleh Particle Swarm Optimization (PSO) dengan memberikan bobot terbaik di setiap atributnya yaitu x0 : 2.67213798 ,x1 : 1.70403491 ,x2 : 1.34646438 ,x3 : 2.82187606 ,x4 : 1.77750849 ,x5 : 1.14161151, x6 : 1.76592477, x7 : 1.48706366, x8 : 0.8921046 dengan hasil menunjukan akurasi sebesar 0.948 atau 95%, MSE 0.0516 atau 0.05% dan AUC 0.9463 atau 95%.

Dengan melakukan optimasi menggunakan metode PSO, maka proses deteksi serangan D-DoS mengalami peningkatan sebesar 3%.

DAFTAR PUSTAKA

Ananto, R. P., Purwanto, Y., & Novianty, A. (2017). Deteksi Jenis Serangan pada Distributed Denial of Service Berbasis Clustering dan Classification Menggunakan Algoritma Minkowski Weighted K-Means dan Decision Tree. *E-Proceeding of Engineering*, 879-885.

Dongoran, A., Rahmadani, S., Zarlis, M., & Zakarias. (2018). Feature Weighting Using Particle Swarm Optimization For Learning Vector Quantization Classifier. *Journal of Physics*, 1-6.

Herditomo, Sunaryo, & Naba, A. (2014). Penerapan Metode Hybrid Fuzzy C-Means dan Particle Swarm

Optimization (FCM - PSO) untuk Segmentasi Citra Geografis. *Jurnal EECCIS*, 27-32.
Juhardi, U., & Andilala. (2019). Optimalisasi Penjualan Motor Menggunakan Algoritma Particle Swarm Optimization (PSO). *Jurnal Media Infotama*.
Mashabi, S. (2021, 09 14). *BSSN: Hingga Agustus 2021 Tercatat 888 Juta Serangan Siber*. Retrieved from Kompas.com: https://nasional.kompas.com/read/2021/09/14/10493771/bssn-hingga-agustus-2021-tercatat-888-juta-serangan-siber#:~:text=JAKARTA%2C%20KOMPAS.com%20%2D%20Kepala,Januari%20hingga%20Agustus%202021%20lalu.
Muhamad, H., Prasojo, C. A., Sugianto, N. A., Surtiningsih, L., & Cholissodin, I. (2017). Optimasi Naive Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 180-184.
Pamungkas, D. P., Utami, E., & Amborowati, A. (2015). Komparasi Pengenalan Citra Tanda Tangan Dengan Metode 2D-PCA dan 2D-LDA. *Citec Journal*, 342-354.
Prasetyo, A., Affandi, L., & Arpandi, D. (2018). Implementasi Metode Naive Bayes Untuk Intrusion Detection System (IDS). *Jurnal Informatika Polinema*, 280-284.
Prasetyo, E. (2012). *Data Mining - Konsep dan Aplikasi Menggunakan MATLAB.* Yogyakarta: ANDI.
Rasila, I., & Ristian, U. (2019). Implementasi Metode Naive Bayes Classifier Pada Sistem Pengklasifikasi Berita Otomatis Berbasis Website (Studi Kasus : Berita Lokal Dari Mediamassa Online Kalimantan Barat). *Jurnal Komputer dan Aplikasi*, 49-60.
Ridho, M. A., & Arman, M. (2020). Analisis Serangan DDoS Menggunakan Metode Jaringan Saraf Tiruan. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 373-379.
Sanmorino, A. (2019). A study for DDOS attack classification method. *International Conference on Advance and Scientific Innovation (ICASI)*, 1-6.
Santosa, B., & Willy, P. (2011). *Particle Swarm Optimization.* Surabaya: Graha Ilmu.
Suniantara, I. P., Suwardika, G., & Soraya, S. (2017). Peningkatan Akurasi Klasifikasi Ketidaktepatan Waktu Kelulusan Mahasiswa Menggunakan Metode Boosting Neural Network. *Jurnal Varian*, 95-102.
Umarani, S., & Sharmila, D. (2014). Predicting Application Layer DDoS Attacks Using Machine Learning Algorithms. *International Scholarly and Scientific Research & Innovation*, 1912-1917.