ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika

ISSN(p): 2338-8323 | ISSN(e): 2459-9638 | Vol. 10 | No. 2 | Halaman 309 - 321
DOI : http://dx.doi.org/10.26760/elkomika.v10i2.309 April 2022

Penerapan Metode Naïve Bayes pada *Honeypot Dionaea* dalam Mendeteksi Serangan *Port Scanning*

DESI KURNIA NURILAHI, RIZAL MUNADI, SYAHRIAL, AL BAHRI

Universitas Syiah Kuala, Indonesia
Email: rizal.munadi@unsyiah.ac.id

Received 12 Agustus 2021 **|** *Revised* 1 September 2021 **|** *Accepted* 4 November 2021

ABSTRAK

Peningkatan serangan terhadap jaringan komputer terus terjadi setiap tahunnya dan dampaknya membuat layanan menjadi terganggu. Pada Penelitian ini Dionaea Honeypot yang merupakan jenis Low Interaction Honeypot, diterapkan untuk mengevaluasi serangan yang terjadi berdasarkan teknik serangan Port Scanning. Data Log yang diperoleh dari pengujian, dianalisis dengan metode Naïve Bayes. Lebih lanjut, data pemetaan Port Scanning dengan menggunakan perangkat lunak Nmap, ditemukan port yang terbuka sebanyak 359 data. Hasil uji klasifikasi dengan menggunakan perangkat lunak WEKA dan penerapan metode Naïve Bayes. Hasil uji klasifikasi diperoleh nilai akurasi sebesar 86,2% dengan nilai rata-rata Precision sebesar 0,885%, Recall sebesar 0,862% dan F-measure sebesar 0,849%. Hasil ini menunjukkan penerapan metode Naïve Bayes berhasil mengklasifikasikan potensi serangan yang dilakukan berdasarkan teknik Port Scanning.

***Kata kunci**: Jaringan Komputer, Low Interaction Honeypot, Port Scanning, Uji Klasifikasi, Akurasi*

ABSTRACT

Increasing attacks on computer networks continue to occur every year, and the impact makes services disrupted. In this study, Dionaea Honeypot, a type of Low Interaction Honeypot, is applied to evaluate attacks based on the Port Scanning attack technique. Log data obtained from the test were analyzed using the Naïve Bayes method. Furthermore, Port Scanning mapping data using Nmap software on the network found 359 open ports data. The results of the classification test using WEKA software and the application of the Naïve Bayes method. The classification test results obtained are accuracy value, 86.2% with an average value of 0.885% Precision, 0.862% Recall and 0.849% F-measure. This result shows that the application of the Naïve Bayes method has succeeded in classifying potential attacks based on the Port Scanning technique.

***Keywords**: Computer Network, Low Interaction Honeypot, Port Scanning, Classification Test, Accuracy*

1. PENDAHULUAN

Serangan dengan motif gangguan dan faktor lainnya kerap dilakukan oleh pihak yang tidak bertanggung jawab terhadap eksistensi jaringan komputer. Berbagai kasus kejahatan berupa serangan terhadap jaringan komputer yang terus berkembang **(Munawar & Putri, 2020)**. Hal ini berdampak pada masalah keamanan sehingga diperlukan perhatian serius agar layanan tidak mengalami gangguan dan dampak kerugian dapat diminimalisasi. Serangan terhadap keamanan jaringan diawali pada tahun 2000 dengan serangan yang dinamakan *Denial of Service* (DOS). Kemudian pada tahun 2004 berkembang menjadi *Distributed Denial of Service* (DDOS) dan enkripsi biner **(Farizy, 2018)**. Tindakan kejahatan siber terus terjadi, termasuk di Indonesia. Berdasarkan pantauan dan hasil rekap Pusat Operasi Keamanan Siber Nasional (PUSOPSKAMSINAS) Badan Siber dan Sandi Negara yang mencatat sebanyak 88.414.296 serangan siber yang telah terjadi sejak 1 Januari hingga 12 April 2020 **(Badan Siber & Sandi Negara, 2020)**. Jenis serangan siber yang rentan terjadi berupa serangan *Port Scanning, Brute Force, Metasploit* dan *Denial Of Service* (DOS) **(Krisna, dkk, 2020).** Fakta menunjukkan bahwa pada penelitian terdahulu telah berhasil mendeteksi serangan *Port Scanning* sebanyak 2053 kejadian, serangan *Brute Force* sebanyak 606 kejadian, serangan *Metasploit* sebanyak 502 kejadian dan serangan DOS sebanyak 428 kejadian **(Romadhan, dkk, 2020)**. Oleh karena itu keamanan jaringan komputer dan *server* harus diperhatikan oleh administrator dengan cara melakukan pencegahan serta identifikasi serangan **(Badan Siber & Sandi Negara, 2020)**.

Berbagai jenis serangan yang terjadi dilakukan dengan tujuan untuk mendapatkan keuntungan dari kerusakan yang ditimbulkan. Target serangan yang berbahaya dilakukan melalui proses mengintai *server* untuk mencari celah sehingga dapat dilancarkannya serangan lanjutan. Salah satu serangan yang melakukan aktivitas mencari informasi ialah serangan *Port Scanning* **(Rohrmann, dkk, 2017)**. Serangan *Port Scanning* memiliki karakteristik untuk mencari informasi *Port* yang terbuka pada suatu jaringan komputer, sehingga dari hasil *Scan port* yang diperoleh, dapat menjadi peluang untuk akses dari jaringan komputer **(Satwika, dkk, 2020), (Valianta, dkk, 2016)**. Maka dari itu serangan *Port Scanning* sangat berbahaya bagi suatu jaringan karena dapat memanfaatkan kelemahan *server* untuk melancarkan aksinya **(Achmad, dkk, 2020)**. Hal ini ditunjukkan pada penelitian yang telah mengimplementasikan *Intrusion Prevention System* (IPS) menggunakan *Snort* dan *Iptables* pada jaringan lokal, yang berhasil mendeteksi serangan *Port Scanning* sebesar 85% sehingga didapatkan kenaikan kinerja CPU sebesar 58,1%. Sebagai akibatnya *server* tidak dapat bekerja dan berhenti melakukan aktivitas atau layanan *down* sementara waktu **(Suwanto, dkk, 2019)**. Untuk menghindari serangan tersebut dibutuhkan sistem keamanan yang dapat mengindentifikasi tindakan awal yang berupa pengintaian atau tindakan pengumpulan informasi dari *server* **(Arman, 2020)**. Beberapa penelitian telah mengusulkan penerapan *Honeypot* sebagai sistem keamanan *server* untuk mengatasi berbagai serangan termasuk serangan *Port Scanning*. Berdasarkan penelitian terdahulu, pengimplementasikan *Low Interaction Honeypot* dengan jenis *Honeyd* telah berhasil menjadi solusi keamanan jaringan dari aktivitas serangan *Denial of Service Attack* dan *Scanning Attack*. Keberhasilan tersebut ditunjukkan dari hasil notifikasi pada rekaman atau *log* ketika serangan masuk **(Fitriana & Khasanah, 2018)**.

Sistem *Honeypot* dirancang sebagai *server* palsu yang kemudian diinstall pada sisi *server* asli yang bertujuan untuk merekam bentuk aktivitas penyusup **(Pandire & Gaikmad, 2018)**. *Honeypot* dirancang untuk mengelabui penyerang yang berusaha masuk dan merusak layanan pada *server* **(Sulaksono & Suharyanto, 2020)**. Berdasarkan interaksinya *Honeypot* dikategorikan dalam tiga bentuk yaitu *High Interaction Honeypot* menerapkan sistem operasi asli sehingga memiliki risiko yang tinggi untuk diterapkan **(Jeremiah, 2019)**, *Medium*

Interaction Honeypot yang dirancang untuk berinteraksi dengan penyerang pada tingkat menengah **(Tripathi & Kumat, 2018)** dan *Low Interaction Honeypot* yang dirancang untuk meniru layanan seperti *server* asli **(Akiyoshi, dkk, 2018)**. Salah satu jenis *Honeypot* yang bersifat *Low Interaction Honeypot* ialah *Dionaea Honeypot* **(Saikawa & Klyuev, 2019)**. *Dionaea Honeypot* menyediakan berbagai layanan seperti FTP, HTTP, SSH, MSSQL, MYSQL, SMB, TFTP, SIP dan lainnya **(Sethia, 2019)**. Dengan kelebihan tersebut *Dionaea* dapat mengelabui penyerang yang ingin mengetahui informasi layanan yang tersedia pada *server*. Penerapan *Dionaea* menghasilkan data *log* yang berupa aktivitas yang dilakukan serangan terhadap *server* **(Ali & Kumar, 2017)**. *Data log* tersebut dapat dianalisis dengan metode yang tepat untuk memperoleh pola serangan sehingga dapat menunjang perbaikan keamanan jaringan. Penelitian yang telah berhasil mengimplementasikan *Honeypot* jenis *Dionaea* untuk menjebak serangan *Port Scanning* sehingga mendapatkan catatan *log* berupa eksploitasi ke MYSQL, layanan *Server message Block* (SMB) dan layanan *Microsoft* RPC **(Cahyanto, 2017)**.

Naïve Bayes merupakan metode pengklasifikasian probabilitas sederhana yang sesuai dengan Teorema Bayes, kemudian dikombinasikan dengan Naïve yang berarti variable *independent* **(Ardyanti, dkk, 2020)**. Naïve Bayes adalah salah satu metode analisis yang dapat mengklasifikasikan pola serangan **(Singh, dkk, 2020)**. Hal ini dibuktikan pada penelitian yang telah mengimplementasikan metode Naïve Bayes untuk mengatasi serangan-serangan baru dan meningkatkan akurasi pendeteksian serangan baru pada *Intrusion Detection System* (IDS). Dari penelitian ini didapatkan hasil klasifikasi dari serangan-serangan baru dengan tingkat akurasi kebenaran yang bagus yaitu sebesar 81-84,67% **(Prasetyo, dkk, 2018)**. Klasifikasi Naïve Bayes dapat diasumsikan ada atau tidaknya suatu ciri tertentu pada sebuah kelas yang tidak berhubungan dengan ciri kelas lainnya **(Adhar, 2021)**. Penerapan metode Naïve Bayes dapat digunakan untuk mendeteksi anomali dengan *Univariate* fitur *Selection*. Dalam penelitian tersebut, hasil penerapan Naïve Bayes dengan fitur yang tidak diseleksi, diperoleh nilai akurasi yang tinggi yaitu 91,43%, sedangkan dengan fitur yang tidak diseleksi diperoleh nilai 91,62%. Oleh karena itu, dapat disimpulkan bahwa tindakan yang dilakukan terhadap data yang ada dapat berpengaruh pada hasil akurasi **(Harianto, ddk, 2021)**.

Keuntungan klasifikasi Naïve Bayes adalah tidak membutuhkan data pelatihan yang cukup besar untuk parameter yang ingin diklasifikasikan **(Haryono, dkk, 2021)**. Salah satu penelitian yang telah menerapkan Naïve Bayes, digunakan untuk pendeteksian serangan jaringan *Local Area Network* (LAN). Hasil penelitian tersebut menunjukkan bahwa perhitungan Naïve Bayes berhasil mengklasifikasikan jenis serangan *Worm* dengan nilai perbandingan sebesar 0,000207899.

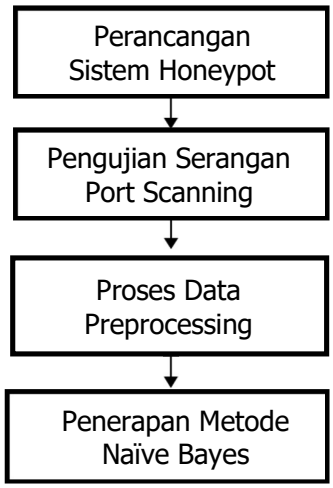
Akurasi merupakan suatu ukuran untuk melihat seberapa baik suatu model dalam mengkorelasikan hasil dengan atribut dalam suatu data. Uji akurasi dilakukan untuk melihat persentase dari total data yang benar dan total data yang salah **(Widiastiwi & Ernawati, 2021)**. Uji akurasi menghasilkan nilai 99-100 persen ataupun dengan pernyataan baik, bagus, kurang baik dan tidak baik. Jika suatu data memiliki tingkat akurasi yang tinggi maka semakin bagus suatu metode klasifikasinya.

Pada penelitian ini, digunakan *Low Interaction Honeypot,* dengan jenis *Honeypot Dionaea*. Tujuan penelitian menggunakan *Honeypot Dionaea* ini dimaksudkan untuk mengevaluasi sistem *Dionaea Honeypot* dalam menangani serangan *Port Scanning* yang dilancarkan. Dengan menggunakan *tools Nmap,* hasil pengujian serangan tersebut dievaluasi berdasarkan *data log* yang berisi aktivitas serangan yang kemudian akan diklasifikasikan dengan metode Naïve Bayes menggunakan perangkat lunak *WEKA*. Hasil analisis yang diperoleh, dapat digunakan sebagai pedoman dalam meningkatkan sistem keamanan jaringan.

2. METODE PENELITIAN

2.1 *Alur Penelitian*

Alur penelitian dijelaskan pada Gambar 1 dengan menggunakan jenis *Low Interaction Honeypot* sebagai pendeteksi serangan *Port Scanning* untuk meningkatkan keamanan jaringan. Adapun proses pengolahan data awal hingga penerapan metode Naïve Bayes dapat dilihat pada gambar alur penelitian berikut:



Gambar. 1 Alur Penelitian

Untuk lebih memahami mengenai alur penelitian dapat dilihat pada penjelasan berikut:

1) Perancangan Sistem *Honeypot*

Pada tahapan ini dilakukan perancangan sistem *Low Interaction Honeypot* dengan melakukan instalasi *Honeypot* jenis *Dionaea* pada *server,* dimana *server* palsu atau *server Honeypot* akan berada di sisi *server* asli. *Honeypot* akan dipasang pada *Ubuntu Server* dengan konfigurasi IP 192.168.11.2. Simulator penyerang dipasang jaringan yang sama menggunakan *Kali Linux* dengan konfigurasi yang berbeda. *Dionaea honeypot* dijalankan pada protokol *Transmission Control Protocol* (TCP) dan *Transport Layer Security* (TLS). Serangan yang disimulasikan terhadap *server* ialah serangan *Port Scanning* yang dicatat dengan menggunakan *tools Nmap*.

2) Pengujian Serangan *Port Scanning*

Tahap ini dilakukan simulasi serangan *Port Scanning* terhadap sistem *Dionaea honeypot* atau *server* palsu yang telah dirancang pada *Ubuntu server*. Pengujian serangan *Port Scanning* dilakukan untuk melacak semua *Port* yang sedang terbuka untuk dilancarkan serangannya. Kemudian *Honeypot* akan merekam dan menyimpan semua bentuk aktivitas ke *log* maupun metode yang diakukan untuk dapat masuk ke *server*.

3) Proses Data *Preprocessing* atau Pemrosesan Awal

Proses ini dilakukan untuk menghilangkan kesalahan yang terjadi secara acak terhadap data mentah maupun data awal. Pada penelitian ini, data *log* serangan terlebih dahulu akan diproses menggunakan *Microsoft Excel*. Proses *Preprocessing* dilakukan terhadap data *log* serangan *Port Scanning.* Data *log* akan dikelompokkan berdasarkan waktu 1 (satu) detik, hal ini dilakukan berdasarkan karakteristik serangan *Port Scanning*. Serangan *Port Scanning* memiliki karakteristik yang dapat melakukan lebih dari tiga serangan dalam waktu 1 (satu) detik. Serangan dilancarkan secara terus-menerus atau *realtime* untuk mendapatkan lebih

banyak informasi mengenai *port-port* yang sedang terbuka, jenis layanan yang sedang dijalankan, versi *server* dan lainnya. Informasi tersebut dimanfaatkan untuk melancarkan aksi serangan terhadap sumber daya *server*. *Data log* yang telah diuraikan *(parsing)* kemudian dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data *testing* merupakan 70% dari data awal, sedangkan data uji atau data *training* merupakan 30% dari data awal.

4) Penerapan Metode Naïve Bayes

Metode Naïve Bayes diterapkan untuk mengklasifikasikan data *log* serangan *Port Scanning*. Naïve Bayes dapat menghitung persentase kinerja sebuah sistem klasifikasi. Sebuah sistem yang baik merupakan dapat mengklasifikasikan data set dengan akurat. Metode Naïve Bayes diterapkan untuk menganalisis *data log* serangan *Port Scanning*, dengan cara menentukan atribut-atribut yang menunjukkan data benar serangan dan bukan serangan. Metode ini menggunakan *prior probability* yang merupakan nilai probabilitas yang diyakini benar sebelum eksperimen dilakukan. Parameter perhitungan Naïve Bayes merupakan tingkat kemiripan yang tertinggi atau disebut metode *maximum likelihood*. Ranah klasifikasi yang dilakukan oleh Naïve Bayes adalah menghitung P(H|X) yang merupakan peluang hipotesis benar atau sahih untuk data sampel X. Berikut persamaan dari Naïve Bayes:

$$P(H|X) = \frac{P(X|H)\,P(H)}{P(X)} \quad (1)$$

Keterangan :
X : Data sampel dengan label yang tidak diketahui
H : Hipotesis bahwa X merupakan data dengan label Y
P(H|X) : Peluang bahwa hipotesis benar untuk data sampel X yang diamati
P(X|H) : Peluang data sampel X, bila diasumsikan hipotesa benar
P(H) : Peluang dari hipotesis H
P(X) : Peluang data sampel yang diamati

Pada penelitian ini, dilakukan uji akurasi untuk melihat tingkat akurasi kebenaran klasifikasi dan kesalahan klasifikasi terhadap *data log* serangan *Port Scanning*. Uji akurasi dilakukan dengan membagi jumlah data yang benar dengan total data uji (benar dan salah). Hasil uji akurasi didapatkan nilai 99-100% dengan pernyataan baik, bagus, kurang baik dan sangat baik. Berikut persamaan untuk menghitung tingkat akurasi:

$$Akurasi = \frac{TP+TN}{TP+FP+TN+F} \times 100\% \quad (2)$$

Keterangan :
TP : *True Positive*
TN : *True Negative*
FP : *False Positive*
FN : *False Negative*

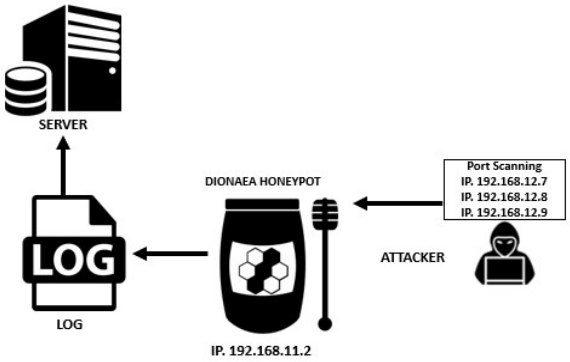
Berdasarkan empat istilah dari nilai di atas, untuk merepresentasikan hasil klasifikasi, dapat dijelaskan sebagai berikut :

a. *True Positive* (TP) merupakan data *positive* yang diprediksikan benar
b. *True Negative* (TN) adalah data *negative* yang diprediksikan benar
c. *False Negative* (FN) adalah data *negative* namun diprediksikan sebagai data *positive*
d. *False Positive* (FP) merupakan data *negative* namun diprediksikan sebagai data *positive*, *False Positive rate* adalah banyaknya data yang mempunyai kelas serangan dan berhasil diklasifikasikan dengan benar berdasarkan penerapan metode Naïve Bayes.

3. HASIL DAN PEMBAHASAN

3.1 Rancangan Sistem

Pada penelitian ini sistem *Honeypot* yang digunakan sebagai upaya untuk mengatasi berbagai serangan jahat terhadap *server*. Jenis *Honeypot* yang diterapkan ialah *Dionaea Honeypot* yang termasuk *Low Interaction Honeypot,* yang bekerja dengan cara menggunakan sistem operasi emulasi ketika berinteraksi dengan penyerang. Pada Gambar 2 dijelaskan serangan terhadap topologi jaringan *Honeypot*, dimana *Honeypot* di-*install* pada sisi *server* yang sesungguhnya *(real)* dengan alamat IP yaitu 192.168.11.2. Kemudian serangan *Port Scanning* dibangkitkan dari IP yang berbeda yaitu 192.168.12.7, 192.168.12.8 dan 192.168.12.9. Ketika serangan *Port Scanning* diluncurkan maka secara otomatis akan terjebak pada *server* palsu yaitu *Honeypot Dionaea*. Semua bentuk aktivitas serangan terhadap *server* akan terekam dan tersimpan pada *Log*. Simulasi penyerangan terhadap *server Honeypot* yang telah disebutkan dapat dilihat seperti Gambar 2.



Gambar 2. Mekanisme Serangan terhadap Honeypot

Pemetaan serangan *Port Scanning* dari alamat IP 192.168.12.7, 192.168.12.8 dan 192.168.12.9 yang dilancarkan terhadap *Dionaea Honeypot*, dipantau dengan menggunakan *tools Nmap*. Serangan *Port Scanning* dijalankan dengan sistem operasi Kali Linux. Serangan *Port Scanning* dengan *tools* dan tipe protokol yang digunakan untuk menyerang *Dionaea Honeypot*, semua informasinya terekam pada *server*. Informasi yang tersimpan merupakan data pelacakan *port-port* yang sedang terbuka atau informasi lainnya dari *server* yang diserang. Beberapa *port* yang terekam seperti *port* 21 (FTP), *port* 80 (HTTP), *port* 3306 (MYSQL) dan *port* 1723 (PPTP) dalam status terbuka. Setelah mendapatkan informasi tersebut penyerang dapat memanfaatkannya untuk melanjutkan serangan dan mematikan sumber daya *server*.

3.2 Hasil Simulasi Serangan *Port Scanning*

Perancangan *server* dilakukan dengan menerapkan sistem operasi *Ubuntu Server* dan *Ubuntu Router* dengan alamat IP 192.168.11.2 sebagai *server.* Kemudian Instalasi *Dionaea Honeypot* dilakukan pada Ubuntu Server dengan IP yang sama dan dimaksudkan sebagai umpan untuk mengelabui penyerang. *Honeypot* yang telah selesai dirancang dengan konfigurasinya akan dilakukan pengujian untuk melihat apakah *tools* berfungsi dengan semestinya. Pengujian dilakukan dengan melancarkan serangan *Port Scanning* terhadap *server Dionaea Honeypot*. Serangan *Port Scanning* dilancarkan menggunakan *Nmap* dengan IP 192.168.12.7, 192.168.12.8 dan 192.168.12.9. Setiap aktivitas *Port Scanning* akan terekam dalam *log Honeypot*. Berdasarkan *data log* serangan *Port Scanning* yang ditampilkan pada Tabel 1 terlihat status serangan, protokol yang diserang, informasi IP *target* dan waktu penyerangan.

Berdasarkan waktu serangan didapat karakteristik serangan yang diperoleh menunjukkan jumlah serangan lebih dari tiga kali dalam waktu 1 detik. Hasil akhir pengujian, diperoleh jumlah serangan *Port Scanning* pada *port* yang terbuka sebanyak 359 data, seperti yang ditunjukkan pada Tabel 1.

Tabel 1. Data Log Serangan *Port Scanning*

Connection Type	*Connection Transport*	*Connection Protocol*	*Local Port*	*Local Host*	*Remote Host*	*Time*
Connect	Tcp	Mirrorc	34477	192.168.11.2	192.168.12.7	27/10/2020 13:43:56
Accept	Tcp	Mirrord	42	192.168.11.2	192.168.12.7	27/10/2020 13:43:56
Accept	Tcp	Ftpd	21	192.168.11.2	192.168.12.7	27/10/2020 13:43:56
Connect	Tcp	Sipsession	5060	192.168.11.2	192.168.12.8	27/10/2020 14:31:36
Accept	Tcp	Sipsession	5060	192.168.11.2	192.168.12.8	27/10/2020 14:31:36
Accept	Tcp	Mysqld	3306	192.168.11.2	192.168.12.8	27/10/2020 14:31:36
Accept	Tcp	Mssqld	1433	192.168.11.2	192.168.12.9	27/10/2020 14:33:04
Accept	Tcp	Blackhole	53	192.168.11.2	192.168.12.9	27/10/2020 14:33:04
Accept	Tcp	Epmapper	135	192.168.11.2	192.168.12.9	27/10/2020 14:33:04
Accept	Tcp	Pptpd	1723	192.168.11.2	192.168.12.9	27/10/2020 14:33:04
Accept	Tcp	Pptpd	1723	192.168.11.2	192.168.12.9	27/10/2020 14:33:04

3.3 Proses *Preprocessing Data*

Proses *preprocessing* data dilakukan sebelum penerapan metode Naïve Bayes. Tahap awal *Preprocessing* dilakukan dengan memilih atribut-atribut yang akan dipergunakan. Hal ini dilakukan karena tidak semua atribut dari data awal dipergunakan seutuhnya. Tahapan ini diperlukan untuk mengurangi derau atau kesalahan dalam menangani data berukuran besar. Sehingga sebelum penerapan metode *data mining* seperti Naïve Bayes perlu dilakukan *preprocessing* data terhadap *data log* serangan. *Preprocessing* dilakukan pada data *log* serangan yang merupakan data rekaman asli dari simulasi serangan *Port Scanning*.

Hasil *data log* yang telah tersimpan pada *Honeypot*, terlebih dahulu dipindahkan ke dokumen pada komputer. Selanjutnya data serangan *Port Scanning* yang tersimpan dalam format *Mysqlite* dikonversi ke dalam format CSV. Proses pengubahan format data dilakukan agar data dapat terbaca oleh Microsoft Excel dan WEKA. *Data log* asli tersebut terlebih dahulu dilakukan proses *PreProcessing* dengan menggunakan Microsoft Excel berdasarkan pendekatan interval waktu. Pengujian serangan *Port Scanning* menghasilkan data sebanyak 359 data. Setelah data dikelompokkan sesuai waktu yaitu 1 detik, maka data menyusut menjadi 96 data. Hal ini dilakukan berdasarkan karakteristik *Port Scanning* yang dapat melakukan *scan* lebih dari tiga *port* dalam waktu 1 detik. Serangan *Port Scanning* dapat melakukan *Scan port* secara terus-menerus atau *realtime*. Pada *data log Port Scanning* setelah di-*parsing* (Tabel 2) dibuktikan bahwa banyaknya *port* yang berhasil disadap oleh penyerang dalam satu waktu. Kemudian metode *parsing data log* serangan *Port Scanning* dapat dilihat pada penjelasan berikut:

a. Jika `connection_type=connect-accept`, `transport=tcp-tls`, `port ≥ 3`, dan `remote host, localhots` yang sama dalam waktu 1 detik maka "`yes`".

b. Jika `connection_type=connect-accept`, `transport=udp-tcp`, `port < 3`, dan `remote host, localhots` yang sama dalam waktu 1 detik maka "`no`".

Data Log hasil pengujian serangan *Port Scanning* (Tabel 2.) menunjukkan bahwa hasil *scan port* yang terdapat pada *Connection Port* yaitu sebanyak 3 *Port* yang dilakukan dalam waktu 1 detik. Sehingga dinyatakan aktivitas tersebut merupakan aktivitas serangan *Port Scanning.* Namun apabila hasil *scan port* pada *Connection Port* kurang dari 3 *port* maka dapat dinyatakan bukan serangan *Port Scanning*. Jenis *Connection Type* yang terdapat pada *data log* berupa *Connect* dan *Accept*. Kemudian *Remote Host* berupa IP Penyerang ataupun IP serangan *Port Scanning* yang terdiri dari tiga IP yaitu 192.168.12.7, 192.168.12.8 dan 192.168.12.9. *Connection Transport* atau protokol komunikasi yang digunakan ialah *Transmission Control Protocol* (TCP). TCP merupakan protokol yang berada pada lapisan *transport* dan bersifat *connection* dan *realiable*. Sedang *Connection Port* atau *port* yang terkoneksi berupa *Port* 445 atau *Server Message Block* (SMB), *Port* 80 atau *Hyper Text Transfer Protokol* (HTTP), *Port* 21 atau *File Transfer Protocol* (FTP) dan lainnya (Tabel 1). Pengujian serangan *Port Scanning* terhadap *Dionaea Honeypot* menghasilkan *data log* seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Data Log *Port Scanning* setelah di-Parsing

Connection Type	***Remote Host***	***Connection Transport***	***Connection Port***	***Time (detik)***	***Port_Scanning***
connect-accept	192.168.12.7	Tcp	3	1 detik	Yes
connect-accept	192.168.12.7	Tcp	9	1 detik	Yes
connect-accept	192.168.12.7	Tcp	22	1 detik	Yes
connect-accept	192.168.12.7	tcp-tls	6	1 detik	Yes
connect-accept	192.168.12.7	Tcp	1	1 detik	No
connect-accept	192.168.12.7	Tcp	1	1 detik	No
connect-accept	192.168.12.7	Tcp	7	1 detik	Yes

Setelah dilakukan pengelompokan data berdasarkan waktu, kemudian dilakukan pengacakan data. Tujuan pengacakan data ialah agar analisis data yang dilakukan sahih, kemudian proses pengacakan data dilakukan. Pada Tabel 2. didapatkan dua pernyataan yaitu "*Yes*" dan "*No*". Pernyataan *Yes* merupakan benar serangan *Port Scanning*, sedangkan *No* adalah bukan serangan *Port Scanning.* Dinyatakan benar serangan *Port Scanning* atau *Yes* karena berdasarkan data yang didapatkan *Remote Host* dengan IP. 192.168.12.7 dapat melakukan *Connection port* lebih atau sama dengan tiga *port* dalam satu waktu yaitu 1 detik dan *Connection Type* berstatus *Connect-Accept*. Sedangkan pernyataan bukan serangan *Port Scanning* atau *No* disebabkan *Romote Host* dengan IP. 192.168.12.7 medapatkan hasil *Connection Port* kurang dari tiga *port* dalam satu waktu yaitu 1 detik, kemudian *Connection Type* berstatus *Connect-Accept*.

3.4 Penerapan Metode Naïve Bayes

Penerapan metode Naïve Bayes dilakukan dengan membuat *data training* sebanyak 70% dan *data testing* sebanyak 30%. Kemudian dari 359 data serangan *Port Scanning* setelah dilakukan proses *Preprocessing* didapatkan data sebanyak 96 data serangan. *Data training* yang merupakan 70% dari *data log* serangan, sehingga didapatkan data training sebanyak 68 data serangan. Selanjutnya *data testing* yang merupakan 30% dari *data log* serangan, maka didapatkan data *testing* sebanyak 28 data sarangan. Pengujian akurasi dilakukan terhadap data *log* serangan *Port Scanning* yang berjumlah 96 data. Namun apabila data serangan lebih dari 96 data maka akan menghasilkan persentase data *testing* dan data *training* yang berbeda, sehingga mengakibatkan hasil akurasi yang berbeda. Begitu juga sebaliknya jika data kurang

dari 96 data maka dapat menghasilkan persentase *data testing*, *data training* dan tingkat akurasi yang berbeda.

Metode Naïve Bayes diterapkan untuk mengklasifikasi data *log* serangan *Port Scanning.* Hal ini dilakukan untuk menunjukkan seberapa banyak yang benar diklasifikasi dan salah diklasifikasikan dari *data log* Serangan *Port Scanning*. Hasil penerapan Naïve Bayes dengan aplikasi WEKA yang terdapat pada Gambar 3. Hasil yang didapat berupa *correctly classified instances* atau tingkat keberhasilan klasifikasi yang baik yaitu 86,2%. Sedangkan *incorrectly classified instances* atau tingkat kesalahan klasifikasi yang kecil yaitu 13,7%.

```
=== Summary ===

Correctly Classified Instances          25               86.2069 %
Incorrectly Classified Instances         4               13.7931 %
Kappa statistic                          0.6329
Mean absolute error                      0.2392
Root mean squared error                  0.3248
Relative absolute error                 56.173  %
Root relative squared error             70.1923 %
Total Number of Instances               29

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.556    0.000    1.000      0.556    0.714      0.680    0.889     0.766     yes
               1.000    0.444    0.833      1.000    0.909      0.680    0.889     0.917     no
Weighted Avg.  0.862    0.307    0.885      0.862    0.849      0.680    0.889     0.870
```

Gambar 3. Uji Akurasi Datalog Port Scanning dengan WEKA

3.5 *Confussion Matrix*

Berdasarkan hasil pengujian yang dilakukan terhadap 68 data serangan *Port Scanning* yang merupakan *data training,* kemudian sebanyak 28 data Serangan *Port Scanning* yang merupakan *data testing. Confusion Matrix* adalah tabel yang menjelaskan jumlah data yang diuji, kemudian dinyatakan benar klasifikasinya dan salah diklasifikasi. Pada Tabel 3 *Confussion Matrix* terdapat dua *Class Label* yaitu *Yes* dan *No*. *Class Label Yes* merupakan hasil prediksi tingkat klasifikasi data yang benar oleh metode Naïve Bayes terhadap data yang benar serangan *Port Scanning*. Sedangkan *Class Label No* adalah hasil prediksi tingkat Klasifikasi data yang benar oleh Naïve Bayes terhadap data yang bukan serangan *Port Scanning.* Pada Tabel 3 ditunjukkan hasil *Confussion Matrix*.

Tabel 3. Confussion Matrix Serangan *Port Scanning*

		Prediksi	
Aktual	***Positive* (Yes)**	***Positive* (Yes)**	***Negative* (No)**
		5	4
	***Negative* (No)**	0	20

Berdasarkan data uji diketahui *matrix* berikut :

TP (*True Positive*) : 5
TN (*True Negative*) :20
FP (*False Positive*) : 4
FN (*False Negative*) : 0

4 Uji Klasifikasi

Uji Klasifikasi dilakukan untuk mengetahui ketelitian atau keakuratan dari hasil klasifikasi. Data yang diuji klasifikasinya merupakan *data log* serangan *Port Scanning*. Kemudian untuk

mengukur tingkat performa metode klasifikasi data maka dilakukan perhitungan *Accuracy*, *Precision, Recall dan F-measure.* Akurasi (*Accuracy*) merupakan suatu ukuran untuk melihat seberapa baik suatu model dalam mengkorelasikan hasil dengan atribut dalam data yang ada. Fungsi akurasi adalah untuk memperkirakan kemungkinan kesalahan dan keberhasilan klasifikasi. Uji akurasi dilakukan untuk mengklasifikasi data, sehingga didapatkan persentase dari total data yang benar. Perhitungan *Precision* merupakan perbandingan jumlah data relevan dari data yang didapat dari jumlah data yang ditentukan. Perhitungan *Precision* dilakukan untuk membagi jumlah data benar positif dan data salah bernilai positif. Sedangkan perhitungan *Recall* dilakukan untuk membagi data benar positif dengan hasil penjumlahan data benar positif dan data salah positif. Kemudian perhitungan *F-measure* merupakan parameter tunggal yang menunjukkan keberhasilan *retrieval.* Pada Tabel 4. diperoleh hasil pengujian *Precision, Recall dan F-measure* terhadap *data log* dari serangan *Port Scanning.* Penelitian uji akurasi klasifikasi ini dilakukan dengan menggunakan aplikasi WEKA dan menerapkan metode *Naïve Bayes*. Berikut didapatkan hasil perhitungan uji akurasi:

Tabel 4. Uji Klasifikasi

Pengujian	Yes	No
Akurasi (%)	86,2%	13,7%
Recall (%)	0,556%	1,000%
Precision (%)	1,000%	0,833%
F-measure (%)	0,714%	0,909%

Pengujian *data log* dilakukan dengan metode Naïve Bayes dan penggunaan aplikasi WEKA. Berdasarkan hasil analisis Naïve Bayes terhadap *data Log* serangan *Port Scanning* yang ditunjukkan pada Tabel 4. menghasilkan nilai *class label Yes* yang dilihat dari *F-Measure* berupa 0,714%. Nilai *class label No* berupa 0,909%, sehingga didapat hasil rata-rata *F-Measure* sebanyak 0,849%. Pengujian *data log* serangan *Port Scanning* dengan metode *Naïve Bayes* pada WEKA menunjukkan tingkat kesalahan klasifikasi yang kecil yaitu 13,7% dibandingkan dengan 86,2% tingkat keberhasilan. Dengan kata lain tingkat akurasi yang didapatkan lebih tinggi dibandingkan *rate error*.

4. KESIMPULAN

Penerapan *Low Interaction Honeypot* serta *Dionaea Honeypot* dalam mendeteksi serangan *Port Scanning* berhasil diterapkan, sehingga didapatkan hasil data *log* yang dianalisis dengan metode Naïve Bayes. Hasil Pengujian serangan *Port Scanning* dengan *tools Nmap* terhadap *Dionaea Honeypot* berhasil ditemukan sebanyak 359 data *scan port* yang terbuka. Sehingga hasil analisis pengujian *data log* serangan dengan metode Naïve Bayes pada perangkat lunak WEKA diperoleh tingkat keberhasilan klasifikasi yang bagus yaitu 86,2%, sedangkan tingkat kesalahan klasifikasi yang kecil yaitu 13,7%. Hasil klasifikasi tersebut membuktikan bahwa penerapan metode Naïve Bayes berhasil mengklasifikasikan *data Log* serangan *Port Scanning*. Berdasarkan data hasil pengujian dengan menggunakan *Dionaea Honeypot*, membuktikan bahwa kinerja *Dionaea Honeypot* cukup baik dalam menangani serangan *Port Scanning.*

DAFTAR RUJUKAN

Achmad, R., Manullang, E. V., & Sanmas, E. R. (2020). Rancang Bangun Aplikasi Deteksi Dan

Penanganan Serangan DDOS Dan Port Scanning Memanfaatkan Snort Pada Jaringan Komputer. *Jurnal Teknologi Informasi*, *8*(1), 2–11.

Akiyoshi, R., Kotani, D., & Okabe, Y. (2018). Detecting Emerging Large-Scale Vulnerability Scanning Activities by Correlating Low Interaction Honeypots with Darknet. *IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, (pp. 658–663).

Ali, P. D. & Kumar, T. G. (2017). "Malware Capturing And Detection In Dionaea Honeypot," *Innovations in Power and Advanced Computing Technologies (i-PACT)*, (pp. 1-5).

Arman, M. (2020). Metode Pertahanan Web Server Terhadap Distributed Slow HTTP DOS Attack. *Jurnal Teknik Informatika Dan Sistem Informasi*, *7* (1), 56–70.

Ardyanti, H., Goejantoro, R., Amijaya, T. D. F. (2020). Perbandingan Metode Klasifikasi Naïve Bayes dan Jaringan Saraf Tiruan (Studi Kasus : PT Asuransi Jiwa Bersama Bumiputera Tahun 2018). *Jurnal Eksponensial, 11* (2), 145-152.

Badan Siber & Sandi Negara (BSSN). (2020, April 20). *Rekap Serangan Siber*. Retrieved from https://bssn.go.id/rekap-serangan-siber-januari-april-2020.

Cahyanto, T. A. (2017). Analisis Dan Implementasi Honeypot Menggunakan Dionaea Sebagai Penunjang Keamanan Jaringan. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699.

Farizy, S. (2018). Mengembangkan Sistem Keamanan Jaringan Komputer Pada Laboratorium Komputer STIMIK Pranata Indonesia Menggunakan Metode Forensik. *Jurnal Teknologi Informasi ESIT*, *14*(02), 55–59.

Fitriana, N., & Khasanah, F. N. (2018). Honeypot Menggunakan Honeyd Sebagai Solusi. *Bina Insani ICT Journal*, *5*(2), 143–152.

Haryono, D., Zulianda, Y., Wirta., Lusiana. (2021). Sistem Pendeteksian Serangan Jaringan Local Area Network (LAN) Menggunakan Algoritma Naïve Bayes. *Jounal Of Information System And Information Engineering (JOISIE), 5*(1), 1-8.

Harianto., Sunyoto, A., Sudarmawan. (2020). Optimasi Algoritma Naïve Bayes Classifier Untuk Mendeteksi Anomaly Dengan Univariate Fitur Selection. *Jurnal Pendidikan Informatika, 4* (2), 40-49.

Jeremiah, J. (2019). Intrusion Detection System to Enhance Network Security Using Raspberry PI Honeypot in Kali Linux. *International Conference on Cybersecurity (ICoCSec)*, (pp. 91-95).

Krisna, I. K., Marta, A., Hartawan, I. N. B., & Satwika, I. K. S. (2020). Analisis Sistem Monitoring Keamanan Server Dengan SMS Alert Berbasis SNORT. *Information System and Emerging Technology Journal*, *1*(1), 25–40.

Munawar, Z., & Putri, N. I. (2020). Keamanan Jaringan Komputer Pada Era Big Data. *J-SIKA| Jurnal Sistem Informasi Karya Anak Bangsa*, *02*(01), 14–20.

Pandire, P. A., & Gaikmad, V. B. (2018). Attack Detection in Cloud Virtual Environment and Prevention using Honeypot. *International Conference on Inventive Research in Computing Applications (ICIRCA)*, (pp. 515–520).

Prasetyo, A., Affandi, L., & Arpandi, D. (2018). Implementasi Metode Naive Bayes Untuk Intrusion Detection System (IDS). *Jurnal Informatika Polinema*, *4*(4), 280.

Rohrmann, R. R., Ercolani, V. J., & Patton, M. W. (2017). Large Scale Port Scanning Through Tor: Using Parallel Nmap Scans to Scan Large Portions of the IPv4 Range. *IEEE International Conference on Intelligence and Security Informatics (ISI)*, (pp. 185–187).

Romadhan, I. A., Syaifudin, S., & Akbi, D. R. (2020). Implementasi Multiple Honeypot Pada Raspberry Pi dan Visualisasi Log Honeypot Menggunakan ELK Stack. *Jurnal Repositor*, *2*(4), 475.

Saikawa, K., & Klyuev, V. (2019). Detection and Classification of Malicious Access using a Dionaea Honeypot. *IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, (pp. 844–848).

Satwika, I. K. S., Sudiarsa, I. W., & Swari, M. H. P. (2020). Intrusion Detection System (IDS) Menggunakan Raspberry Pi 3 Berbasis Snort Studi Kasus : STMIK Stikom Indonesia. *STMIK Journal*, *XV*, (pp. 2–7).

Sethia, V., & A, Jeyasekar. (2019). Malware Capturing and Analysis using Dionaea Honeypot. *International Carnahan Conference on Security Technology (ICCST)*, (pp. 1–4).

Singh, S., Kumari, K., Gupta, S., Dua, A., & Kumar, N. (2020). Detecting Different Attack Instances of DDOS Vulnerabilities on Edge Network of Fog Computing using Gaussian Naive Bayesian Classifier. *IEEE International Conference on Communications Workshops (ICC Workshops)*, (pp. 1-6).

Sulaksono, W. A., & Suharyanto, C. E. (2020). Implementasi Honeypot Sebagai Sistem Keamanan Jaringan Pada Virtual Private Server. *Jurnal Nasional Informatika Dan Teknologi Jaringan*, *1*.

Suwanto, R., Ruslianto, I., & Diponegoro, M. (2019). Implementasi Intrusion Prevention System (IPS) Menggunakan Snort Dan IPTable Pada Monitoring Jaringan Lokal Berbasis Website. *Jurnal Komputer Dan Aplikasi*, *7*(1), 97–107.

Tripathi, S., & Kumar, R. (2018). Raspberry Pi as an Intrusion Detection System, a Honeypot and a Packet Analyzer. *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, (pp. 80–85).

Valianta, S. A., Salim, T., & Stiawan, D. (2016). Identifikasi Serangan Port Scanning Dengan Metode String Matching. *Annual Research Seminar (ARS) Fakultas Ilmu Komputer UNSRI*, (pp. 466–471).

Widiastiwi, Y., Ernawati, I. (2021). Klasifikasi Penyakit Batu Ginjal Menggunakan Algoritma Decision Tree C4.5 Dengan Membandingkan Hasil Uji Akurasi. *Jurnal IKRA-ITH Informatika, 5*(2), 128-135.