IMPLEMENTASI *NAIVE BAYES* DAN KNN PADA DETEKSI SERANGAN DDoS PADA JARINGAN METRO

TUGAS AKHIR

MUCHAMAD OKTARIN JATMIKA
41518210001

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022

IMPLEMENTASI *NAIVE BAYES* DAN KNN PADA DETEKSI SERANGAN DDoS PADA JARINGAN METRO

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

Oleh:

MUCHAMAD OKTARIN JATMIKA

41518210001

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2022

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518210001

Nama : **Muchamad Oktarin Jatmika**

Judul Tugas Akhir : Implementasi Naïve Bayes dan K-NN pada Deteksi Serangan DDOS pada Jaringan Metro

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 18 Juli 2022

SEPULUH RIBU RUPIAH 10000 METERAI TEMPEL 2F8AJX907431791

Muchamad Oktarin Jatmika

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Muchamad Oktarin Jatmika
NIM	:	41518210001
Judul Tugas Akhir	:	Implementasi Naïve Bayes dan K-NN pada Deteksi Serangan DDOS pada Jaringan Metro

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Juli 2022

METERAI TEMPEL
A2772AJX907431786

Muchamad Oktarin Jatmika

Universitas Mercu Buana

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Muchamad Oktarin Jatmika
NIM	:	41518210001
Judul Tugas Akhir	:	Implementasi Naïve Bayes dan K-NN pada Deteksi Serangan DDOS pada Jaringan Metro

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis			Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi			Diajukan	✓
		Jurnal Nasional Terakreditasi				
		Jurnal International Tidak Bereputasi			Diterima	
		Jurnal International Bereputasi		✓		
	Disubmit/dipublikasikan di :	Nama Jurnal	: 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI 2022)			
		ISSN				
		Link Jurnal	: https://edas.info/index.php?c=28547			
		Link File Jurnal Jika Sudah di Publish	:			

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui Dosen Pembimbing TA	Jakarta, 18 Juli 2022
	METERAI TEMPEL 5ADB6AJX907431781
Rahmat Budiarto, Dr. Prof	Muchamad Oktarin Jatmika

Universitas Mercu Buana

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

Nama Mahasiswa	:	Muchamad Oktarin Jatmika
NIM	:	41518210001
Judul Tugas Akhir	:	Implementasi Naïve Bayes dan K-NN pada Deteksi Serangan DDOS pada Jaringan Metro

Tugas Akhir ini telah diperiksa dan disetujui

Jakarta, 18 Juli 2022

Menyetujui,

(Rahmat Budiarto, Dr. Prof

Dosen Pembimbing

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518210001
Nama	:	Muchamad Oktarin Jatmika
Judul Tugas Akhir	:	Implementasi Naïve Bayes dan K-NN pada Deteksi Serangan DDOS pada Jaringan Metro

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022

Runi
(Runi)

Saruni Dwiasnati, ST, MM, M.Kom

Penguji 1

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518210001
Nama	:	Muchamad Oktarin Jatmika
Judul Tugas Akhir	:	Implementasi *Naive Bayes* Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022

Drs. Acmad Kodar, MT

Penguji 2

LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41518210001
Nama	:	Muchamad Oktarin Jatmika
Judul Tugas Akhir	:	Implementasi *Naive Bayes* Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022

Sabar Rudiarto, S.Kom., M.Kom

Penguji 3

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM	:	41518210001
Nama	:	Muchamad Oktarin Jatmika
Judul Tugas Akhir	:	Implementasi Naive Bayes Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022

Menyetujui,

(Rahmat Budiarto, Dr. Prof)
Dosen Pembimbing

Mengetahui,

(Wawan Gunawan, S.Kom, MT)
Koord. Tugas Akhir Teknik Informatika

(Ir. Emil R. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan Allah SWT atas rahmat dan karunia-Nya sehingga Tugas Akhir yang berjudul " Implementasi *Naive Bayes* Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro" dapat diselesaikan dalam jangka waktu yang sudah ditentukan. Laporan tugas akhir ini dibuat sebagai syarat untuk LULUS sebagai sarjana Ilmu Komputer dari Universitas Mercu Buana.

Penulis menyadari bahwa Laporan Tugas Akhir ini masih jauh dari kata sempurrna dikarenakan keterbatasan kemampuan dan pengetahuan yang penulis dapatkan. Dalam menyelesaikan Laporan Tugas Akhir ini, penulis mendapat banyak bantuan, bimbingan, dan dukungan Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua, Ayah, Ibu, Kakak dan Adik, yang tidak pernah lelah medukung dan memberi semangat agar saya bisa menyelesaikan kuliah dengan baik serta tepat waktu, juga tak pernah luput mendoakan yang terbaik untuk proses meraih gelar sarjana bagi saya.
2. Prof, Dr Rahmat Budiarto selaku Dosen Pembimbing Tugas Akhir yang telah memberikan masukan saat bimbingan dan meluangkan sebagian besar waktunya untuk melakukan bimbingan dalam penyusunan tugas akhir ini hingga selesai.
3. Bapak Emil R. Kaburuan, Ph.D. selaku Kepala Program Studi Teknik Informatika Universitas Mercu Buana.
4. Seluruh Dosen Program Studi Teknik Informatika yang sudah memberikan ilmu yang bermanfaat selama kuliah berlangsung. Memberi kesempatan untuk belajar, berkarya, dan juga berkembang.
5. Seluruh Staff Administrasi dan Tata Usaha yang telah banyak membantu dan memberikan kemudahan terima kasih atas semua pelayanan dan arahannya.
6. Semua pihak dan personal yang tidak dapat disebutkan satu per satu yang terlibat dalam pembuatan Tugas Akhir ini sehingga dapat selesai dengan baik.

Akhir kata, penulis berharap segala kekurangan penulisan, eksperimen dan cara penjelasan dapat dimaafkan. Untuk itu, kritik dan saran pembaca sangat dihargai dan diharapkan semoga Tugas Akhir ini dapat memberikan manfaat bagi para pembaca.

Jakarta, 07 Juli 2022

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

NASKAH JURNAL

Detection of DDOS Attacks on Metro Network using Naïve Bayes and KNN

Muchamad Oktarin Jatmika
Dept. of Informatics
Faculty of Computer Science
Mercu Buana University
Jakarta, Indonesia
41518210001@student.mercubuana.ac.id

Wawan Gunawan
Dept. of Informatics
Faculty of Computer Science
Mercu Buana University
Jakarta, Indonesia
wawan.gunawan@mercubuana.ac.id

Rahmat Budiarto
Dept. of Informatics
Faculty of Computer Science
Mercu Buana University
Jakarta, Indonesia
rahmat.budiarto@mercubuana.ac.id

Deris Stiawan
Dept. of Informatics
Faculty of Computer Science
Sriwijaya University
Palembang, Indonesia
deris@unsri.ac.id D

Abstract— Broad increase in data consumption in society and industry trigger network operators looking to upgrade their metro networks with higher bandwidth requirements. Service providers and operators are challenged to find a simple, the most efficient and cost-effective way of meeting the demand with new speeds and standards on the horizon. Distributed Denial of Service (DDoS) attack is a cyber attack that uses a method to flood internet network traffic on the server, system, or network of the targeted attack. The occurrence of DDoS attacks on the metro networks can make the operating system unable to operate properly and even crash. DDoS can be prevented by monitoring traffic regularly, increasing server resource capacity and implementing multiple protection strategies. This paper implements DDoS attacks detection system by combining Information Gain feature Selection and Naïve Bayes classifier. As comparison, K-Nearest Neighbor (KNN) classifier is also considered. The main aim is to improve the detection accuracy as such may help the metro network optimally provides the necessary bandwidth. Experimental results using CICIDS-2018 dataset show that the KNN outperforms Naïve Bayes classifier with the accuracy level 99%

Keywords— Metro Network, DDoS attack, Naïve Bayes, KNN

I. Introduction

Broad increase in data consumption in society and industry trigger network operators looking to upgrade their metro networks with higher bandwidth requirements. Service providers and operators are challenged to find a simple, the most efficient and cost-effective way of meeting the demand with new speeds and standards on the horizon.

Distributed Denial of Service (DDoS) attack is a cyber attack that uses a method to flood internet network traffic on the server, system, or network of the targeted attack. The occurrence of DDoS attacks on the metro networks can make the operating system unable to operate properly and even crash. DDoS can be prevented by monitoring traffic regularly, increasing server resource capacity and implementing multiple protection strategies.

By referring to research work by Susanto and Jatikusumo [1], this paper implements DDoS attacks detection system by combining Information Gain feature Selection and Naïve Bayes classifier. As comparison, K-Nearest Neighbor (KNN) classifier is also considered. The main aim is to improve the detection accuracy as such may help the metro network optimally provides the necessary bandwidth.

II. Literature Review

Table 1 summarizes previous works related to this paper.

TABLE I. SUMMARY OF RELATED WORKS

Ref. #	Method	Summary
Sugianti et al.[2], 2020	Sugeno Fuzzy	The significant features are number of users, number of packets, packet length. Experiments using Sugena Fuzzy on MATLAB provide accuracy level of 90% on HTTP-based DDoS attack.
Sihombing et al. [3], 2019	SVM Classifier	Detection system for DDoS attack on SDN architecture. The features are taken from flow entries. The system classifies either the traffic is normal or attack. The detection system achieves accuracy detection up to 96.83% and average detection time is 67.80 ms. The system is also able to reduce the attack traffic sent to the victim hosts.
Harto & Basuki [4], 2021	Random Forest	The authors reveal the Random Forest algorithm works well in detecting the DDoS attack, with accuracy level of 90% and average detection time was 0.3 seconds. Processing time for taking decision was good enough, i.e.: 281 ms. The generated decision tree was 15 trees and does not affect the engine workload.
Sukarno & Nugroho [5], 2019	KNN and Decision Tree (DT)	DT algorithm performs better than KNN and in term of running-time, DT algorithm is also better than KNN. DT algorithm achieves 99.91% accuracy, while KNN achieves 98.94% in detecting DDoS attack.
Riadi et al. [6], 2019	Naïve Bayes & SVM	Naïve Bayes has probability value between 0.1 to 0.8 in term of Radviz and graph distribution, while SVM provides higher accuracy values.
Aziz et al. [7], 2019	Artificial Neural Network (ANN)	The authors conclude that accuracy of attacks detection using signature-based IDS should be reviewed by considering statistical approaches. The ANN-based IDS provides accuracy level of 95.2381%. The ANN method also can be utilized for digital forensics investigation.
Purba et al. [8], 2022	Deep Q-Network (DQN), Support Vector Regression (SVR) & Logistic Regression (LR)	CICDDoS2019 dataset is used. The proposed DQN is able to detect 11 DDoS attack types and benign/normal data with a better accuracy value than LR and SVR algorithms. DQN achieves up to 96% accuracy level.
Nasution and Basuki [9], 2021	C5.0 algorithm	CICDDoS2019 dataset is used. The dataset contains 56279 instances including 25133 instances DDoS attacks traffic and 31146 instances of normal traffic. The accuracy, precision and recall of C5.0 algrithm is 98.38%, 98.39%, and 98.37%, respectively and the processing time is 16.84 seconds.
Farid et al. [10], 2021	Fuzzy Mamdani	Experiments are carried out using MATLAB. QoS during scenario with attack and without attack is measured. The highest throughput during the attack scenario was 5456 bps with 30 nodes, while during the normal scenario was 26247 bps with 30 nodes. On the delivery time, for the attack scenario was 98.478 ms with 10 nodes, while for the normal scenario was 5.53 ms with 20 nodes. The highest Recall value on Cooja was 98.62% with 20 nodes.
Azis, Azhar dan Saifuddin [11], 2020	KNN	The authors propose the use of machine learning algorithm on the RYU controller to deal with the DDoS attacks, such as SYN Flooding attack. The experiment uses linear topology on Mininet that generates .Pcap format files. Thus, the average number of packets that coming in and out, and the successful of performing mitigation against suspicious DDoS attacks can be measured.
Doshi et al.	KNN, Random Forests, Decision Trees,	The five classifiers are implemented on dataset constructed from consumers devices of

[12], 2018	SVM, and Deep Neural Networks	IoT network testbed and provide accuracy above 99%. This initial result motivates researchers to investigate the use of machine learning algorithms for anomaly detection to protect IoT systems.
Dong and Sarem [13], 2019	KNN, Naïve Bayes, SVM, DDoS Detection Algorithm based on the Degree of Attack (DDADA), DDoS Detection Algorithm based on Machine Learning (DDAML), Cognitive-Inspired Computing Support Vector Machine (CIC-SVM)	Experimental results show that the proposed DDAML outperforms the other algorithms. The proposed DDADA dan DDAML are suitable to be implemented in real SDN environment.
Kacha-vimath et al. [14], 2020	Naïve Bayes and KNN	A DDoS attack detection system for enterprise network was proposed. The proposed system consists of three functions, i.e.: preprocessing, machine learning engine and performance evaluation. KNN algorithm performs better the Naïve Bayes algorithm. The authors suggest considering variant of deep learning architecture to be incorporated into the proposed system as well as more significant features of the DDoS attack traffic to increase the detection accuracy.
Reddy and Thilagam [15], 2018	Naïve Bayes Classifier	The authors carry out DDoS attack experiments on NS2 simulator to measure the network performance. Simulation results show the proposed approach was able to reduce the effect of DDoS attacks. The network running the proposed mechanism identifies 80% of the legitimate traffic, while The network without running the proposed mechanism was not able to identify legitimate traffic in unfriendly environment.
Chena et al. [16], 2018	Random Forest	The authors introduce a new method to reduce the DDoS attack traffic on the TLD server. Traffic filtering based on machine learning algorithm is implemented in the main recursive DNS server of the Internet. Experimental results show the FPR value of 0 and FNR value of 4.36%, that mean the accuracy and the required performance in practice is fulfilled.

III. RESEARCH METHOD

Figure 1 shows the proposed research method. Process in data science is constructed by three stages, i.e.: data collection, data transformation and data analysis [17]. This paper uses Canadian Institute for Cybersecurity (CICDS) 2018, consists of1000 data, divided into 2, i.e.: testing data (200) and training data (800). The second stage is data preprocessing where the raw data, shown in Table 2 is transformed into form that easy to understand using Min Max Scaled sklearn.

TABLE II. EXAMPLES OF DATA

No	src_ port	dst_ port	flow_ duration	tot_ fwd _pkts	tot_ bwd _pkts	Label
1	37882	80	8660	1	1	ddos
2	80	63287	5829	4	3	ddos
3	63095	80	3396	1	1	ddos
4	52341	80	2390	1	1	ddos
5	80	57459	17362	4	3	ddos
6	80	56276	201316	4	3	ddos
7	55330	53	22123	2	2	Benign
8	53799	443	3095495	4	2	Benign
9	56889	3389	1127340	8	7	Benign
10	51263	443	105120546	16	24	Benign

Firstly, the maximum and minimum values of the data are determined (refer to Table 3). Then the Max Min Scaled is performed using (1). Next, is calculating the standard (std) value using (2). Table 4 shows the scaled data.

$$x_{Std} = \frac{x - Min(x)}{Max(x) - Min\,(x)} \quad (1)$$

$$x_{Scaled} = x_{Std} * (Max - Min) + Min \quad (2)$$

In the third stage, feature selection is conducted using information gain method where the relevancy to the DDoS attack traffic is measured. From an initial experiment, we obtain the probability of the attack traffic is 0.6, while for benign traffic is 0.4. The probability for calculating the gain can be performed using (3). The calculation of the probability is shown in Table5.

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \qquad (3)$$

Then the values of Information Gain is calculated using (4). Information Gain calculation results are displayed in Table 6.

$$IG(c,t) = S(c) + \sum_{j\epsilon value(t)}^{m} \frac{Cj}{C} S(cj) \qquad (4)$$

The fourth stage is data classification using Naïve Bayes as well as KNN classifiers. The last stage is validation using confusion matrix.

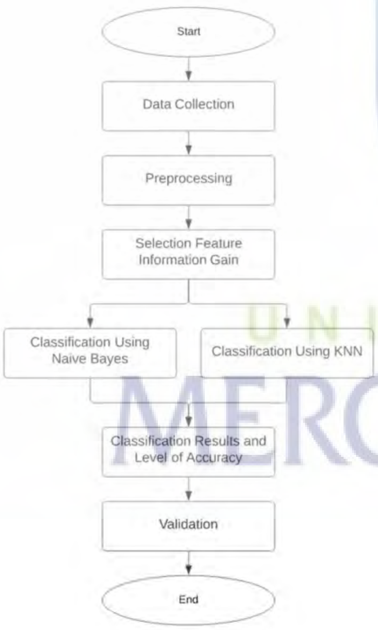

Figure 1. The proposed research method

TABLE III. MAX AND MIN DATA

No	src_port	dst_port	flow_duration	tot_fwd_pkts	tot_bwd_pkts
1	37882	80	8660	1	1
2	80	63287	5829	4	3
3	63095	80	3396	1	1
4	52341	80	2390	1	1
5	80	57459	17362	4	3
6	80	56276	201316	4	3
7	55330	53	22123	2	2
8	53799	443	3095495	4	2
9	56889	3389	1127340	8	7
10	51263	443	105120546	16	24
MAX	**63095**	**63287**	**105120546**	**16**	**24**
MIN	**80**	**53**	**2390**	**1**	**1**

TABLE IV. THE SCLAED DATA

No	src_port	dst_port	flow_duration	tot_fwd_pkts	tot_bwd_pkts
1	-1909707383	-268112159	-2175672521	-14	-22
2	-403295999	-212100178	-14644336178	-59	-68
3	-3180745139	-268112159	-85318520608	-14	-22
4	-2638614491	-268112159	-60044541888	-14	-22
5	-403295999	-192568207	-43618968044	-59	-68
6	-403295999	-188603498	-50577100396	-59	-68
7	-2789295959	-177624305	-55580142267	-29	-45
8	-2712115187	-148467108	-77768861587	-59	-45
9	-2867888267	-113579013	-28322432574	-119	-160
10	-2584270355	-148467108	-26409686308	-239	-551

TABLE V. PROBABILITY

No	src_port	dst_port	flow_duration	tot_fwd_pkts	tot_bwd_pkts
1	-314306006	-43842240	-3624603664605	-2.33	-3.67
2	-66375800	-34683048	-2439701473555	-9.83	-11.33
3	-523497637	-43842240	-1421380374711	-2.33	-3.67
4	-434271968	-43842240	-1000323644157	-2.33	-3.67
5	-66375800	-31489140	-7266786238439	-9.83	-11.33
6	-66375800	-30840823	-8425989738380	-9.83	-11.33
7	-68860744	-43568226	-1388922174458	-7.25	-11.25
8	-66955343	-3641646	-1943408057869	-14.75	-11.25
9	-70800991	-27859003	-7077645546056	-29.75	-40
10	-63799174	-3641646	-6599659057568	-59.75	-137.75

TABLE VI. INFORMATION GAIN RESULTS

No	src_port	flow_duration	tot_fwd	tot_bwd

			_pkts	_pkts
1	-50925530240	-580179339198506	-3.73	-5.85
2	-107545600	-390515631430496	-15.73	-18.13
3	-84819870400	-227516054955904	-3.73	-5.87
4	-70363053120	-160118778370026	-3.73	-5.87
5	-107545600	-116317248119682	-15.73	-18.13
6	-107545600	-134872267725273	-15.73	-18.13
7	-83678878800	-16674042680398	-8.70	-13.50
8	-81363455640	-233306584762277	-17.70	-13.50
9	-86036648040	-849672977232737	-35.70	-48
10	-77528110680	-792290589246819	-71.70	-165.30

A. Naïve Bayes Classifier

Naïve Bayes classifier starts with inputting the information gain data, followed by calculating the Gaussian distribution values, then the classification itself. The flowchart for the classifier is shown in Figure 2.

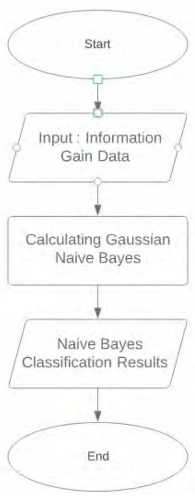

Figure 2. Naïve Bayes classifier flowchart

Next step is to separate the dataset based on the class (DDoS and Benign), as shown in Table 7 and Table 8.

TABLE VII. DDOS CLASS DATASET

No	src_port	dst_port	flow_ duration	tot_ fwd _pkt	tot_b wd _pkt
1	-50925530240	-71496576	-5.80179E+14	-3.73	-5.87
2	-107545600	-56560047566	-3.90516E+14	-15.73	-18.13
3	-84819870400	-71496576	-2.27516E+14	-3.73	-5.87
4	-70363053120	-71496576	-1.60119E+14	-3.73	-5.87
5	-107545600	-51351522005	-1.16317E+15	-15.73	-18.13
6	-107545600	-50294266387	-1.34872E+16	-15.73	-18.13

TABLE VIII. BENIGN CLASS DATASET

No	src_port	dst_port	flow_ duration	tot_ fwd_ pkt	tot_ bwd_ pkt
7	-836788788	-53287291	-1.6674E+15	-8.7	-13.5
8	-8136345564	-445401325	-2.33307E+17	-17.7	-13.5
9	-8603664804	-3407370413	-8.49673E+16	-35.7	-48

TABLE IX. MEAN & STD-DEV. VALUES OF EACH CLASS

Class	src_port	dst_port	flow_ duration	tot_ fwd_ pkts	tot_ bwd_ pkts
Mean					
ddos	-344051810	-2640338761	-2.66812E+15	-9.73	-12
Benign	-83692994	-130201967	-1.06647E+17	-20.7	-25
StDev.					
Ddos	3908072362628923046		5.31247E+15	6.5726	6.71507
Benign	2336628176	18337978	1.17332E+17	13.7477	19.9185

Final step in Naïve Bayes algorithm is to determine the Gaussian Naïve Bayes using (5)

$$P(xi|C) = \frac{1}{\sqrt{2\mu a^2 c,i}} \exp\left(-\frac{(xi-\mu c,i)^2}{2a^2 c,i}\right) \quad (5)$$

B. KNN Classifier

Flowchart for KNN classifier is shown in Figure 3. The Euclidian distance is calculated using (6).

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \qquad (6)$$

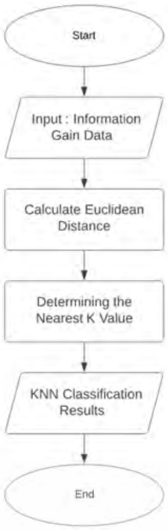

Figure 3. KNN classifier flowchart

IV. Experimental Results and Discussion

A. Naïve Bayes Classifier

The Gaussian Naïve Bayes calculation results for the 10th testing data are shown in Table 10. The Gaussian value falls under the Class ddos.

TABLE X . GAUSSIAN NAÏVE BAYES CALCULATION

	src_port	dst_port	flow_ duration	tot_fwd _pkts	tot_bwd_ pkts	Label
10th Data (Testing)	-7752811068	-445401325	-7.92291E+18	-71.7	-165.3	ddos
ddos	1.09812E-06	1.56852E-06	0	7.74034E-21	1.0379E-114	-1.94553E+42
Benign	2.54195E-07	8.35528E-06	0	0.000110547	1.50658E-12	0

B. KNN Classifier

TABLE XI. EUCLIDEAN DISTANCE CALCULATION

No	src_port	dst_port	flow_duration	tot_fwd_pkts	tot_bwd_pkts	Jarak (Euclidean Distance)	Rank	Label
1	-50925530240	-71496576	-5.80179E+14	-3.73	-5.87	7.92233E+18	4	ddos
2	-107545600	-56560047566	-3.90516E+14	-15.73	-18.13	7.92252E+18	3	ddos
3	-84819870400	-71496576	-2.27516E+14	-3.73	-5.87	7.92268E+18	2	ddos
4	-70363053120	-71496576	-1.60119E+14	-3.73	-5.87	7.92275E+18	1	ddos
5	-107545600	-51351522005	-1.16317E+15	-15.73	-18.13	7.92174E+18	5	ddos

Based on the calculation of the Euclidean distance for the 10th data with value K=1, the classification provides labels with the highest probability for ddos label.

C. Validation using Confusion Matrix

The confusion matrix for Naïve Bayes classifier is shown in Table 12 and Figure 4.

TABLE XII. NAÏVE BAYES CLASSIFIER CONFUSION MATRIX

Actual	Prediction	
	Positif (True)	**Negatif (False)**
Positif	34	68
Negatif	0	98

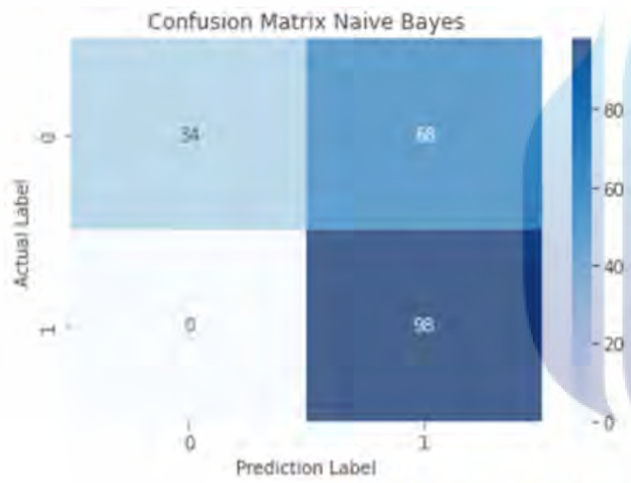

Figure 4. Confusion matrix of Naïve Bayes classifier

Benign Class

Accuracy level of Benign class is calculated using (7)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \quad (7)$$

$$Accuracy = \frac{34+98}{34+98+0+68} * 100\% = 66\%$$

Precision of class Benign class is calculated using (8)

$$Precision = \frac{TP}{TP+FP} * 100\% \quad (8)$$

$$Precision = \frac{98}{0+98} * 100\% = 100\%$$

Recall of Benign class is calculated using (9).

$$Recall = \frac{TP}{TP+FN} * 100\% \quad (9)$$

$$Recall = \frac{34}{34+68} * 100\% = 33\%$$

F1-Score of Benign class is calculated using (10).

$$F1\ Score = \frac{(2*Recall*Precision)}{(Recall+Precision)} * 100\% \quad (10)$$

$$F1\ Score = \frac{(2*0{,}33*1)}{(0{,}33+1)} * 100\% = 50\%$$

DDoS Class

$$Precision = \frac{98}{68+98} * 100\%$$

$$Recall = \frac{98}{0+98} * 100\% = 100\%$$

Confusion matrix of KNN classifier is shown in Figure 5. and Table 13.

TABLE XIII. NAÏVE BAYES CLASSIFIER CONFUSION MATRIX

Actual	Prediction	
	Positif (True)	**Negatif (False)**
Positif	101	1
Negatif	0	98

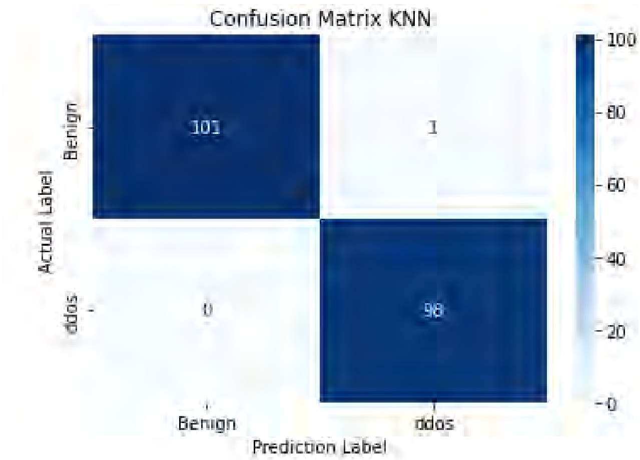

Figure 5. Confusion matrix of KNN classifier

Benign Class

$$Precision = \frac{98}{0 + 98} * 100\% = 100\%$$

$$Recall = \frac{101}{101 + 1} * 100\% = 99\%$$

$$F1\ Score = \frac{(2 * 0{,}99 * 1)}{(0{,}99 + 1)} * 100\% = 100\%$$

DDoS Class

$$Precision = \frac{98}{1 + 98} * 100\%$$

$$Recall = \frac{98}{0 + 98} * 100\%$$

$$F1\ Score = \frac{(2 * 1 * 0{,}99)}{(1 + 0{,}99)} * 100\% = 99\%$$

V. Conclusion

This paper has discussed a DDoS detection system for metro networks. The system utilizes two intelligent classifiers. Overall, KNN classifier outperforms Naïve Bayes classifier. Naïve Bayes classifier provides only 66% accuracy level, while KNN provides 99% accuracy level. For future work, the authors plan to carry out research on other feature selection methods and combined with deep learning classifiers for improving the accuracy of the detection system.

References

[1] S.K Setianto and D. Jatikusumo, "Employee Turnover Analysis Using Comparison of Decision Tree and Naïve Bayes Prediction Algorithms on K-Means Clustering Algorithms at PT. AT," *Jurnal Mantik,* vol. 4, no. 3, pp. 1573-1581, 2020.

[2] N Sugianti, Y. Galuh, S. Fatia and K. F. H. Holle, "Deteksi Serangan Distributed Denial of Services (DDOS) Berbasis HTTP Menggunakan Metode Fuzzy Sugeno," *JISKa,* vol. 4, no. 3, pp. 18-26, 2020.

[3] J. C. J. Sihombing, D. P. Kartikasari and A. Bhawiyuga, "Implementasi Sistem Deteksi dan Mitigasi Serangan Distributed Denial of Service (DDoS) menggunakanSVM Classifier pada Arsitektur Software-Defined Network (SDN)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* vol. 3, no. 10, pp. 9608-9613, 2019.

[4] M. K. Harto and A. Basuki, "Deteksi Serangan DDoS Pada Jaringan Berbasis SDN Dengan Klasifikasi Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* , vol. 5, no. 4, pp. 1329-1333, 2021.

[5] I. Ramadhan, P. Sukarno and M. A. Nugroho, "Analisis Perbandingan Algoritma K-Nearest Neighbor dan Decision Tree Dalam Mendeteksi Distributed Denial of Service," *e-Proceeding of Engineering,* vol. 6, no. 2, pp. 8548-8558, 2019.

[6] I. Riadi, R. Umar and F. D. Aini(3), "Analisis Perbandingan Detection Traffic Anomaly Analisis Perbandingan Detection Traffic Anomaly Machine (SVM)," *ILKOM Jurnal Ilmiah,* vol. 11, no. 1, pp. 17-22, 2019.

[7] M. Aziz, R. Umar and F. Ridho, "Implemetasi Jaringan Saraf Tiruan Untuk Mendeteksi Serangan DDoS Pada Forensik Jaringan," *QUERY: Jurnal Sistem Informasi* , vol. 3, no. 1, pp. 46-52, 2019.

[8] R. Purba, W. S. Lestari and M. Ulina, "Deteksi Serangan DDoS Mengunakan Deep Q-Network," *Jurnal Teknik Informatika dan Sistem Informasi* , vol. 9, no. 1, pp. 648-658, 2022.

[9] E. O. Nasution and A. Basuki, "Implementasi Algoritme C5.0 Untuk Klasifikasi Serangan DDoS," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* vol. 5, no. 1, pp. 389-395 , 2021.

[10] M. Farid, I. Wahidah and A. I. Irawan, "Analisis Pendeteksian Serangan Denial Of Service (DOS) Menggunakan Logika Fuzzy Metode Mamdani Pada Jaringan Internet of Things (IOT)," *e-Proceeding of Engineering* , vol. 8, no. 1, pp. 121-128, 2021.

[11] M. M. Azis, Y. Azhar and Saifuddin, "Analisa Sistem Identifikasi DDoS Menggunakan KNN Pada Jaringan Software Defined Network (SDN)," *REPOSITOR,* vol. 2, no. 7, pp. 915-922, 2020.

[12] R. Doshi, N. Apthorpe and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," *IEEE Security and Privacy Workshops,* pp. 29-35, 2018.

[13] S. Dong and M. Sarem, "DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Improved KNN With the Degree of DDoS," *IEEE Access,* p. 5039–5048, 2019.

[14] A. V. Kachavimath, S. V. Nazare and S. S. Akki, "Distributed Denial of Service Attack Detection using Naïve Bayes and K-Nearest Neighbor for Network Forensics," *International Conference on Innovative Mechanisms for Industry Applications (ICIMIA).,* pp. 711-717, 2020.

[15] K. G. Reddy and P. S. Thilagam, "Naïve Bayes Classifier to Mitigate the DDoS Attacks Severity in Ad-Hoc Networks," *International Journal of Communication Networks and Information Security (IJCNIS),* vol. 12, no. 2, pp. 221-226, 2020.

[16] L. Chena, Y. Zhang, Q. Zhao, G. Gen and Z. Yan, "Detection of DNS DDoS Attacks with Random Forest Algorithm on Spark," *International Workshop on Big Data and Networks Technologies,* p. 310–315, 2018.

[17] D. Dedy and A. Cherid,"Data Mining Pengolahan Data Calon Pekerja Migran Indonesia (PMI) Dengan Penerapan Metode Klustering K-Means dan Metode Klasifikasi K-Nearest Neighbor (KNN): Studi Kasus PT. Sam," *Jurnal Ilmiah Teknik Informatika,* vol. 9, no. 2, pp. 166-182, 2020.

KERTAS KERJA

Ringkasan

Pada bagian Literature Review ini ditampilkan hasil review terhadap beberapa literatur atau jurnal ilmiah yang terkait dengan penelitian ini yaitu Implementasi *Naive Bayes* Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro. Literature Review ini terdiri dari 10 Artikel jurnal umum Nasional dan 5 jurnal Internasional.

Analisis dan Perancangan ditampilkan analisis permasalahan yaitu diperlukan sebuah sistem klasifikasi yang mampu mendeteksi jenis serangan pada traffic layer Mikorotik OS terhadap Distributed Denial of Services (DDoS).

Source Code berisi kumpulan kode-kode bahas pemprogaman python. Source Code ini dijadikan dalam satu folder Bernama lampiran Source Code.

Dataset berisi data yang diambil dari data Canadian Institute for Cybersecurity CICDS 2018 dengan jumlah 1000 data. Yang nantinya akan digunakan dalam penelitian Implementasi *Naive Bayes* Dan KNN Pada Deteksi Serangan DDOS Pada Jaringan Metro.

Tahapan Eksperimen merupakan penjelasan tahapan dari eksperimen yang telah dilakukan penulis

Hasil Eksperiman merupakan isi semua eksperimen menggunakan beberapa klasifikasi algoritma. Sesuai dengan metode maupun jenis klasifikasi yang digunakan pada penelitian ini yaitu klasifikasi algoritma Naïve Bayes dan KNN.