Deteksi Serangan DDoS Mengunakan Deep Q-Network

Ronsen Purba*[1], Wulan Sri Lestari[2], Mustika Ulina[3]

[1,2,3]Teknologi Informasi, Fakultas Informatika, Universitas Mikroskil
Jl. Thamrin, No. 124,140,112 Medan
e-mail: ***[1]ronsen@mikroskil.ac.id**, [2]wulan.lestari@mikroskil.ac.id,
[3]mustika.ulina@mikroskil.ac.id

Abstrak

Distributed Denial of Service Attack (DDoS) merupakan serangan dengan mengkompilasi beberapa sistem di internet dengan zombie/agen yang terinfeksi dan membentuk jaringan botnet. Serangan DDoS mengakibatkan kerugian finansial, hilangnya produktivitas, kerusakan merek, penurunan peringkat kredit dan asuransi serta terganggunya hubungan pelanggan, dan pemasok. Selain itu, teknologi IoT juga rentan terhadap serangan DDoS berskala besar. Untuk mencegah terjadinya serangan DDoS maka dibutuhkan model yang dapat mendeteksi adanya serangan DDoS. Pada penelitian ini, kami mengusulkan Deep Q-Network (DQN) untuk mendeteksi serangan DDoS. DQN merupakan algoritme reinforcement learning yang menggabungkan deep learning dan Q-learning. Penerapan DQN digunakan untuk meningkatkan akurasi deteksi serangan pada dataset. Pada penelitian ini, dataset yang digunakan untuk mendeteksi adanya serangan DDoS atau tidak adalah CICDDoS2019 dataset yang disediakan oleh Canadian Institute for Cybersecurity. Berdasarkan perbandingan metode yang dilakukan didapatkan hasil metode DQN yang diusulkan dapat mendeteksi 11 serangan DDoS dan benign/normal data dengan nilai akurasi yang lebih baik dibandingan metode LR dan SVR. Hasil penelitian menunjukkan model yang diusulkan memiliki nilai akurasi 96% dan lebih baik dibandingkan metode LR dan SVR.

Kata kunci— Deep Q-Network, DDoS, Deteksi Serangan.

Abstract

Distributed Denial of Service Attack (DDoS) is an attack by compiling multiple systems on the internet with infected zombies/agents and forming a network of botnets. DDoS attacks resulted in financial losses, lost productivity, brand damage, downgrades of credit and insurance ratings, and disrupted customer and supplier relationships. In addition, IoT technology is also vulnerable to large-scale DDoS attacks. To prevent DDOS attacks, a model that can detect DDoS attacks is needed. In this research, we propose Deep Q-Network (DQN) to detect DDoS attacks. DQN is a reinforcement learning algorithm that combines deep learning and q-learning. The application of DQN is used to improve the accuracy of attack detection on the dataset. In this paper, the dataset used to detect DDoS attacks or not is the CICDDoS2019 dataset provided by the Canadian Institute for Cybersecurity. Based on the comparison of the methods carried out, the results of the proposed DQN method can detect 11 DDoS attacks and benign/normal data with better accuracy values compared to the LR and SVR methods. The results showed that the proposed model had an accuracy value of 96% and was better than LR and SCR methods.

Keywords— Deep Q-Network, DDoS, Attack Detection.

1. PENDAHULUAN

Beberapa tahun terakhir serangan *Distributed Denial-of-Service* (DDoS) telah menyebabkan kerugian yang signifikan untuk pengguna *Internet of Thing* (IoT) dalam bidang industri, dan pemerintahan [1]. *Cybercriminal* mengetahui teknologi IoT lebih rentan terhadap serangan DDoS skala besar karena keterbatasan sumber daya. Adanya serangan DDoS dapat menyulitkan organisasi seperti terganggunya aktivitas online, data *cloud* organisasi, dan perangkat yang terhubung dengan internet [2]. Serangan DDoS bertujuan untuk membuat suatu *server* memiliki jumlah permintaan yang terlalu tinggi sehingga server tidak dapat ditangani [3].

Pemantauan alamat IP adalah salah satu teknik yang digunakan untuk melawan serangan DDoS [4]. Namun serangan DDoS terutama pada lapisan aplikasi sangat sulit dideteksi karena pola permintaan serangan sama dengan pola permintaan yang sah membuat sistem pertahanan tradisional tidak berjalan [5]. Untuk mengatasi permasalahan tersebut, beberapa peneliti [6][7][8][9][10] menggunakan *machine learning* untuk mendeteksi *cyberattack*. Hasil penelitian tersebut menunjukkan *machine learning* mampu mendeteksi serangan dengan baik dibandingkan teknik *data mining*. Pada *machine learning*, *dataset* sangat penting untuk melatih dan menguji model deteksi [11]. Beberapa peneliti [12][13][14][15][16] mencoba mengembangkan *dataset* DDoS untuk tujuan deteksi *cyberattack*. Namun terdapat masalah seperti lalu lintas tidak lengkap, data anonim, dan skenario serangan yang sudah ketinggalan zaman yang membatasi pengujian dan validasi model deteksi yang diusulkan [17].

Peneliti [17] menyusun *dataset* DDoS komprehensif yang disebut CICDDoS2019 yang terdiri dari 80 fitur lalu lintas jaringan, 1 kelas normal(*benign*) dan 12 kelas DDoS yang dibuat menggunakan CICFlowMeter. Pada penelitian tersebut dilakukan deteksi DDoS dengan membandingkan 3 metode yaitu *Naïve Bayes, Random Forest* (RF) dan ID3. Hasil dari penelitian tersebut metode ID3 memiliki nilai presisi 78% dan lebih baik dibandingkan metode lainnya [17]. Tantangan dalam mendeteksi metode menggunakan *dataset* CICDDoS2019 yaitu *imbalanced* data. *Dataset* CICDDoS memiliki 50006249 data serangan DDoS dan 56863 data *benign*/normal, *imbalanced* data dapat menyulitkan model *machine learning* untuk mendeteksi *cyberattack* [18].

Pada makalah ini, untuk meningkatkan akurasi deteksi serangan DDoS pada CICDDoS2019 diusulkan model *Deep Q-Network* (DQN). DQN merupakan metode gabungan *Q-Learning* dan *Deep Learning*. Beberapa penelitian menggunakan DQN untuk mendeteksi *Intrusion Detection Systems* (IDS) [8][19] dan mengklasifikasi pada *imbalance data* teks dan citra [20]. Kontribusi pada penelitian ini adalah menghasilkan model dengan akurasi tinggi untuk mendeteksi ada atau tidaknya serangan DDoS menggunakan DQN.

2. METODE PENELITIAN

Penelitian ini fokus pada masalah bagaimana menghasilkan model yang tepat untuk mendeteksi adanya serangan DDoS atau tidak. Berikut tahapan-tahapan dalam mendeteksi serangan DDoS:

2.1 Pengumpulan Data

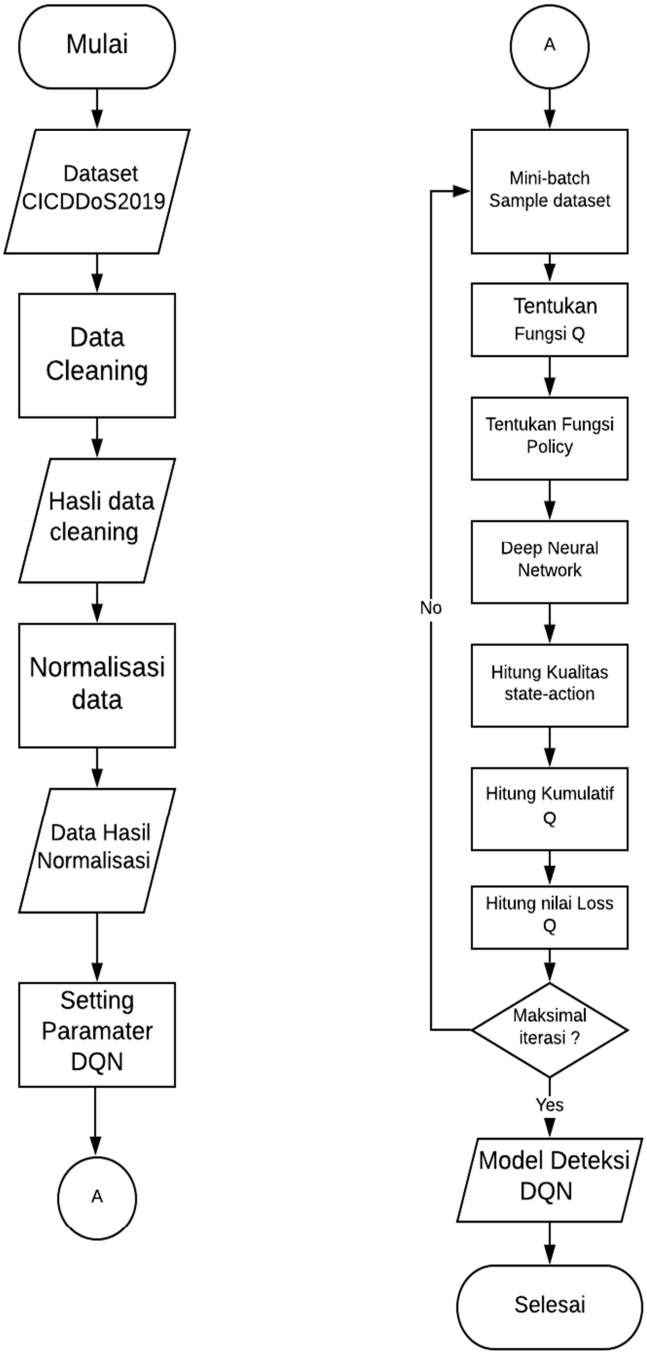
Data yang digunakan pada penelitian ini adalah CICDDOS2019 *dataset* yang disediakan oleh Canadian Institute for Cybersecurity. Pada penelitian ini menggunakan 24 fitur terbaik dari 80 fitur yang ada. Pemilihan 24 fitur tersebut berdasarkan hasil dari penelitian Sharafaldin *et al* [17]. Dataset dibagi menjadi 2 bagian yaitu:

1. *Training Data* merupakan data yang akan digunakan untuk melatih model DQN dalam mendeteksi serangan DDoS. *Training data* terdiri dari 12 serangan DDoS dan 1 *Benign*/normal. Pada makalah ini menggunakan *dataset* serangandengan jumlah 40394 data serangan sedangkan pada data *Benign*/normal memiliki jumlah 8798 data.
2. *Testing Data* merupakan data yang akan digunakan untuk menguji model DQN yang dihasilkan dari proses training model. *Testing data* memiliki 6400 data serangan dan 880 data *Benign*/normal.

2. 2 Implementasi Metode DQN

Pada tahap ini dilakukan proses analisis dan desain terhadap metode DQN untuk diimplementasikan pada sistem deteksi serangan DDoS yang dapat digunakan untuk membedakan terjadi serangan atau tidak. Berikut tahapan-tahapan implementasi DQN:

1. *Preprocessing Data*
   Pada tahap ini dibagi dalam 2 bagian yaitu *cleaning data* dan normalisasi data. Pada proses *cleaning* data bertujuan mengecek dan menghilangkan nilai NaN dan *Infinity* yang terdapat pada *dataset*. Pada proses normalisasi data dilakukan penskalaan data 0 sampai 1 untuk mempermudahkan pemprosesan data pada model.
2. Tahap *Training*
   Tahap *training* merupakan proses yang bertujuan untuk melatih model DQN dalam mengenali *dataset* dan membentuk sebuah model berdasarkan pelatihan tersebut. Gambar 1 merupakan flowchart dari tahap *training*.



Gambar 1. *Flowchart* Tahap *Training*

Berikut ini penjelasan tahap *training*.

a. Dilakukan pemberian label kelas pada *dataset* dengan memberi label "1" untuk setiap data serangan DDoS dan label "0" untuk data *Benign*/Normal.
b. Kemudian dilakukannya *prepocessing data* menggunakan library pandas python.
c. Dilakukannya inisialisasi parameter DQN yang akan digunakan. Tabel I menunjukkan nilai parameter dan arsitektur DQN yang diterapkan pada makalah ini.

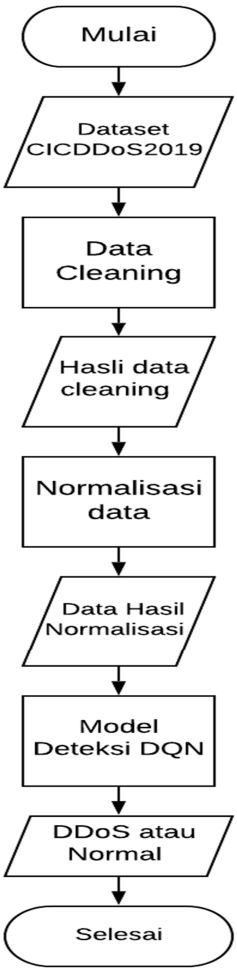
d. Kemudian menentukan *mini-batch sample* dataset sebanyak *batch_size* yang sudah ditentukan.
e. Menentukan fungsi Q dan setting maksimal *reward.*
f. Menentukan fungsi *Policy* pada fungsi Q saat ini.
g. Menerapkannya *Deep Neural Network* (DNN). Pada makalah ini DNN dibangun menggunakan TensorFlow *python* dengan 3 *hidden layers* dan aktivasi ReLU untuk semua lapisan dengan *hidden layer.* ReLu berfungsi untuk memastikan nilai fungsi Q bernilai positif. Menerapkan *dropout* pada DNN di setiap *hidden layer* untuk menghindari terjadinya *overfitting*. Untuk menghitung nilai *loss* pada proses training DQN digunakan *Mean Square Error* (MSE).

Nilai parameter yang digunakan untuk membangun model DQN pada tahap *training* dapat dilihat pada Tabel 1.

Tabel 1. Nilai Parameter Model DQN

Parameter	**Nilai**
Episode	60
Discount Factor (γ)	0.01
Batch	25
Learning Rate	0.001
Epoch	145
Optimizer	Adam
Jumlah *Neuron Hidden Layer* 1	264
Jumlah *Neuron Hidden Layer* 2	128
Jumlah *Neuron Hidden Layer* 3	32
Dropout	0.01

3. Tahap *Testing*
Tahap *testing* merupakan proses memanggil model yang sudah dibangun pada tahap *training* seperti terlihat pada Gambar 2.



Gambar 2. *Flowchart* Tahap *Testing*

Berikut ini penjelasan tahap *testing*.

a. Memasukkan *dataset testing*.
b. Melakukan proses *cleaning* data yang bertujuan mengecek dan menghilangkan nilai NaN yang terdapat pada *dataset*.
c. Melakukan proses normalisasi data untuk penskalaan data 0 sampai 1 untuk mempermudahkan pemprosesan data pada model.
d. Selanjutnya memanggil model deteksi DQN untuk mendeteksi *data testing* merupakan data serangan DdoS atau tidak. Setelah proses deteksi selesai dilakukan perhitungan Akurasi dengan *Confusion Matrix, F-measure* dan *Precision*.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan beberapa pengujian untuk mengetahui akurasi model DQN yang dibangun untuk mendeteksi serangan DDoS pada dataset CICDDoS2019.

3.1. Pengaruh Parameter Discount Factor (γ) terhadap akurasi

Untuk mengetahui pengaruh *Discount Factor* (γ) terhadap akurasi model DQN maka pada pengujian ini akan digunakan data dengan rincian yaitu 49237 *data training* dan 7280 *data testing*. Hasil pengujian pengaruh *Discount Factor* (γ) dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian Pengaruh *Discount Factor*

Nilai *Discount Factor* (λ)	**Akurasi (%)**	**F1-*Score* (%)**	***Precision* (%)**
0.001	**95,58**	**97,68**	96.70
0.01	**95,58**	97,50	**96.98**
0.99	87,90	93,56	87,90

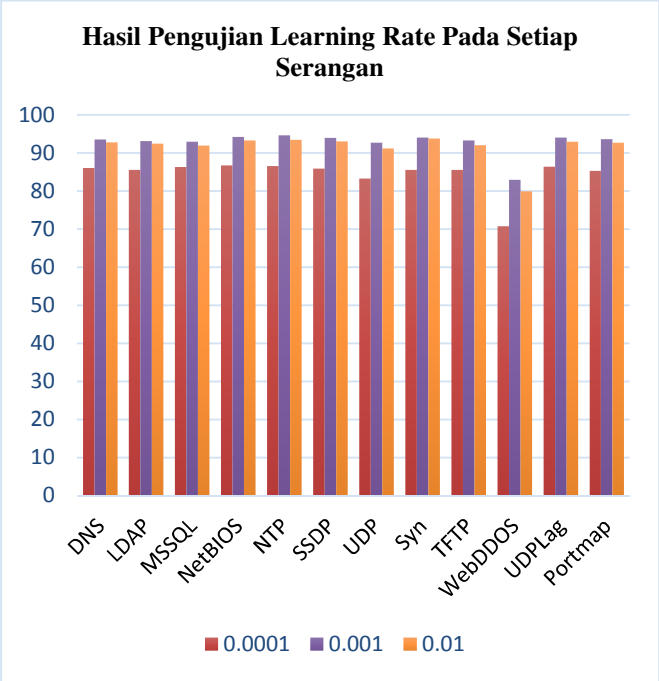
Berdasarkan hasil pada Table 2, dapat dilihat bahwa nilai *Discount Factor* (λ) memiliki pengaruh yang signifikan dalam mendeteksi serangan DDoS. Nilai *Discount Factor* (λ) yang lebih kecil memiliki hasil akurasi yang lebih baik. Dengan menerapkan nilai kecil pada λ memberikan model mempelajari lebih dalam pada *reward* saat ini dibandingkan fokus pada *reward* di masa mendatang. Hal ini dapat diterima mengingat bahwa keadaan selanjutnya (*next state*) tidak berkorelasi dengan keadaan saat ini (*state*) dan tujuan model sebenarnya untuk membuat prediksi yang benar untuk keadaan saat ini.

3.2. Pengaruh Parameter Learning Rate Terhadap Akurasi

Untuk mengetahui pengaruh *Learning Rate* terhadap akurasi model DQN maka dilakukan pengujian dengan nilai 3 learning yang berbeda. Pengujian dilakukan menggunakan *dataset* dengan rincian data sebagai berikut:

1. 1 *dataset training* memiliki 49237 data yang terdiri dari 12 Serangan DDoS dan 8798 data *Benign*/normal.
2. 12 *dataset testing* yang merupakan 12 serangan DDoS. Pada 11 *dataset* serangan masing-masing *dataset* memiliki 1000 data serangan dan 1000 data *benign*/normal. Sedangkan pada *dataset* serangan WebDDoS memiliki jumlah 200 data serangan dan 1800 data *benign*/normal.

Hasil pengujian dapat dilihat pada Gambar 3.



Gambar 3. Grafik Hasil Pengujian Learning Rate

Gambar 3 menunjukkan nilai parameter *learning rate* memiliki pengaruh terhadap hasil akurasi yang didapat. Nilai *learning rate* 0.001 memiliki nilai akurasi paling tinggi dengan rata-rata 93.35% pada 11 serangan. Pada serangan WebDDoS, DQN model menghasilkan nilai akurasi 83%. Hal ini terjadi dikarenakan kelas serangan *WebDDoS* merupakan *imbalanced dataset*. Dari hasil pengujian dapat disimpulkan jika *learning rate* di set terlalu rendah, pelatihan akan berkembang sangat lambat karena membuat pembaruan yang sangat kecil pada bobot di jaringan. Namun, jika *learning rate* di set terlalu tinggi, ini dapat menyebabkan perilaku berbeda yang tidak diinginkan dalam fungsi *loss.*

3.3. Perbandingan Metode dalam Deteksi pada Imbalanced Data

Pengujian ketiga dilakukan dengan membandingkan metode *Support Vector Regression* (SVR), *Logistic Regression* (LR) dan DQN untuk melihat model yang memiliki akurasi terbaik dalam deteksi serangan DDoS pada *imbalanced data.* Pengujian ini menggunakan data *training* dengan jumlah *400* data Webddos dan 8798 data *Benign*/normal dan data *testing* sebanyak 200 data serangan WebDDoS dan 1800 data *Benign*/normal. Tabel 3 menunjukkan hasil pengujian perbandingan metode dalam deteksi serangan DDoS pada *imbalanced data.*

Tabel 3. Hasil Perbandingan Deteksi Serangan DDoS Pada *Imbalanced Data*

Metode	**Akurasi (%)**	**F1-*Score* (%)**	***Precision* (%)**
SVR	57.17	39,50	25.80
LR	58.50	40.30	26.50
DQN (Metode yang diusulkan)	**83.00**	**54.46**	**49.19**

Pada Tabel 3, dapat dilihat bahwa metode yang diusulkan (DQN) menghasilkan nilai akurasi yang lebih baik dibandingkan metode lainnya yaitu 83%. Namun dapat dilihat nilai *F1-Score* dan *Precision* yang dihasilkan masih rendah yaitu dibawah 60%.

3.4. Perbandingan Metode Deteksi DDoS

Pengujian perbandingan metode DQN dengan 2 metode deteksi lainnya yaitu *Support Vector Regression* (SVR) dan *Logistic Regression* (LR) untuk melihat model yang memiliki akurasi terbaik dalam mendeteksi DDoS. Pada pengujian ini menggunakan 6400 data 11 serangan DDoS dan 880 data *Benign*/normal. Tabel 4 menunjukkan hasil pengujian perbandingan metode deteksi.

Tabel 4. Perbandingan Metode Deteksi DDoS

Metode	**Akurasi (%)**	**F1-*Score* (%)**	***Precision* (%)**
SVR	92.72	96.00	92.80
LR	92,95	96.10	93.00
DQN (Metode yang diusulkan)	**96,26**	**97.88**	**97.13**

Tabel 4 menunjukkan metode DQN yang diusulkan pada makalah ini memiliki nilai akurasi yang lebih baik dibandingkan 2 metode deteksi lainnya. Metode yang diusulkan berhasil mendeteksi serangan DDoS dengan akurasi 96,26%.

4. KESIMPULAN

Berdasarkan hasil pengujian yang dihasilkan maka nilai parameter *discount factor* dan *learning rate* sangat berpengaruh pada akurasi deteksi serangan. Nilai *discount factor* yang lebih kecil dapat membuat metode DQN fokus pada nilai *reward* data saat ini sehingga dapat menghasilkan nilai akurasi yang baik. Selain itu model DQN yang diusulkan dapat mendeteksi adanya 11 jenis serangan DDoS atau tidak dengan nilai akurasi 96,26%. Namun, untuk mendeteksi serangan pada *imbalanced data* model DQN hanya menghasilkan akurasi 83%.

5. SARAN

Saran untuk penelitian selanjutnya adalah perlunya meningkatkan akurasi pada *imbalanced data* untuk mendeteksi adanya serangan DDoS sehingga dapat memberikan dampak yang lebih baik untuk diterapkan pada aplikasi keamanan data. Salah satu metode yang dapat digunakan untuk deteksi serangan pada *imbalanced data* adalah *Deep AutoEncode*-DQN.

UCAPAN TERIMA KASIH

Terima kasih kepada Universitas Mikroskil yang sudah mendanai dan mendukung penelitian ini melalui skema hibah kompetisi Internal Kreativitas dan Inovasi Dosen.

DAFTAR PUSTAKA

[1] Li, Q., Meng, L., Zhang, Y., dan Yan, J., 2019, *DDoS Attacks Detection Using Machine Learning Algorithms, Commun. Comput. Inf. Sci.*, Vol. 1009, hal 205–216.

[2] Mishra, A., Sharma, S., dan Pandey, A., 2020, *An Enhanced DDoS TCP Flood Attack Defence System in a Cloud Computing, SSRN Electron. J.*

[3] Li, J., 2020, *Detection Of Ddos Attacks Based On Dense Neural Networks*, *Autoencoders And Pearson Correlation Coefficient (Halifax: Dalhousie University)*, Hal 89.

[4] Maciá-Fernández, G., Rodríguez-Gómez, R. A., dan Díaz-Verdejo, J. E., 2010, *Defense Techniques For Low-Rate Dos Attacks Against Application Servers, Comput. Networks*, vol. 54, no. 15, hal 2711–2727.

[5] Sharma, V., Verma, V., dan Sharma, A., 2019, *Detection of DDoS Attacks Using Machine Learning in Cloud Computing, Commun. Comput. Inf. Sci.*, Vol. 1076, hal 260–273.

[6] Agrawal S., dan Agrawal, J., 2015, *Survey On Anomaly Detection Using Data Mining Techniques, Procedia Comput. Sci.*, Vol. 60, No. 1, Hal 708–713.

[7] Buczak, A. L., dan Guven, E., 2016, *A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, IEEE Commun. Surv. Tutorials*, vol. 18, No. 2, hal 1153–1176.

[8] Mishra, P., Varadharajan, V., Tupakula, U., dan Pilli, E. S., 2019, *A Detailed Investigation and Analysis of Using Machine Learning Techniques For Intrusion Detection, IEEE Commun. Surv. Tutorials*, Vol. 21, No. 1, hal 686–728.

[9] Rawat, S., Srinivasan, A., dan R, V., 2019, *Intrusion Detection Systems Using Classical Machine Learning Techniques Versus Integrated Unsupervised Feature Learning and Deep Neural Network*, *https://arxiv.org/abs/1910.01114*.

[10] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., dan Venkatraman, S., 2019, *Deep Learning Approach for Intelligent Intrusion Detection System*, *IEEE Access*, Vol. 7, hal 41525–41550.

[11] Brownlee, J., *Data Learning and Modeling*, https://machinelearningmastery.com/data-learning-and-modeling/, Diases Pada Tanggal 1 Oktober 2020.

[12] Subbulakshmi, T., Balakrishnan, K., Shalinie, S. M., Anandkumar, D., Ganapathisubramanian, V., dan Kannathal, K., 2011, *Detection of DDoS Attacks Using Enhanced Support Vector Machines With Real Time Generated Dataset, 3rd Int. Conf. Adv. Comput. ICoAC 2011*, hal 17–22.

[13] Prasad, K. M., Reddy, A. R. M., and Rao, K. V., 2014, *DoS and DDoS Attacks: Defense, Detection and TracebackMechanisms -A Survey*, *Glob. J. Comput. Sci. Technol.*, Vol. 14, No. 7.

[14] Brown, C., Cowperthwaite, A., Hijazi, A., dan Somayaji, A., 2009, *Analysis of The 1999 DARPA/Lincoln Laboratory IDS Evaluation Data with NetADHICT*, *IEEE Symp. Comput. Intell. Secur. Def. Appl. CISDA 2009*, no. Cisda.

[15] Singh, K. J., dan De, T., 2015, *An Approach of Ddos Attack Detection Using Classifiers*, *Emerging Research in Computing, Information*, Communication and Applications.

[16] Yu, S., Zhou, W., Jia, W., Guo, S., Xiang, Y., dan Tang, F., 2012, *Discriminating DDoS Attacks From Flash Crowds Using Flow Correlation Coefficient, IEEE Trans. Parallel Distrib. Syst.*, Vol. 23, No. 6, hal 1073–1080.

[17] Sharafaldin, I., Lashkari, A. H., Hakak, S., dan Ghorbani, A. A., 2019, *Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy, Proc. - Int. Carnahan Conf. Secur. Technol.*, Vol. 2019-Octob.

[18] Sun, Y., Wong, A. K. C., dan Kamel, M. S. , 2009, *Classification of imbalanced data: A Review, Int. J. Pattern Recofgnit. Artif. Intell.*, Vol. 23, No. 4, hal 687–719.

[19] Sethi, K., Sai Rupesh, E., Kumar, R., Bera, P., dan Venu Madhav, Y., 2020, *A Context-Aware Robust Intrusion Detection System: A Reinforcement Learning-Based Approach," Int. J. Inf. Secur.*, Vol. 19, No. 6, hal 657–678.

[20] Lin, E., Chen, Q. dan Qi, X., 2020, *Deep Reinforcement Learning For Imbalanced Classification, Appl. Intell.*