IMPLEMENTASI ALGORITMA *RANDOM FOREST* DAN SELEKSI FITUR *CHI-SQUARE* PADA KLASIFIKASI SERANGAN *BOTNET* DI JARINGAN *INTERNET OF THINGS* (IoT)

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

OLEH :

Muhammad Arun Nugraha

09011181823003

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2022

LEMBAR PENGESAHAN

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DAN SELEKSI FITUR *CHI-SQUARE* PADA KLASIFIKASI SERANGAN *BOTNET* DI JARINGAN *INTERNET OF THINGS* (IoT)

TUGAS AKHIR

Program Studi Sistem Komputer
Jenjang S1

Oleh

Muhammad Arun Nugraha
09011181823003

Indralaya, Juni 2022

Mengetahui,

Ketua Jurusan Sistem Komputer	**Pembimbing Tugas Akhir**
Dr. Ir. H. Sukemi, M.T.	**Deris Stiawan, MT., Ph.D., IPU.**
NIP. 196612032006041001	**NIP. 197806172006041002**

HALAMAN PERSETUJUAN

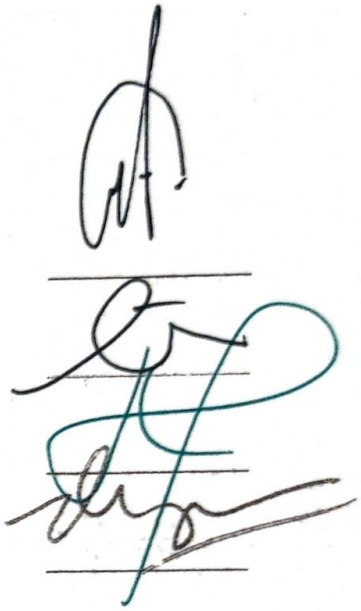
Telah diuji dan lulus pada :

Hari : Rabu

Tanggal : 20 Juli 2022

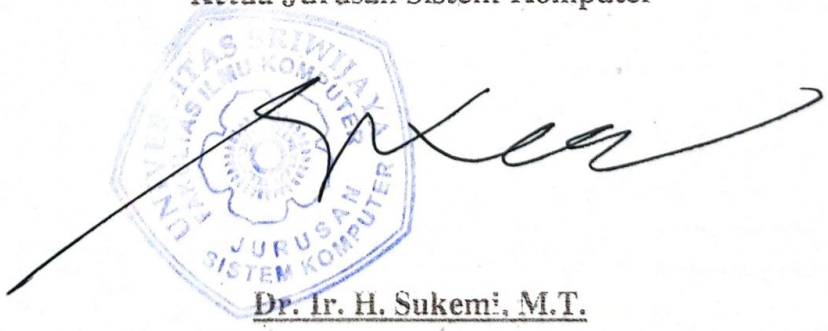
Tim Penguji :

1. **Ketua Sidang** : **Ahmad Zarkasi, M.T.**
2. **Sekretaris Sidang** : **Tri Wanda Septian, M.Sc.**
3. **Penguji Sidang** : **Huda Ubaya, M.T.**
4. **Pembimbing** : **Deris Stiawan, M.T., Ph.D., IPU**



Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Muhammad Arun Nugraha

NIM : 09011181823003

Judul : Implementasi Algoritma *Random Forest* dan Seleksi Fitur *Chi-Square* pada Klasifikasi Serangan *Botnet* di Jaringan *Internet of Things* (IoT)

Hasil Pengecekan Software iThenticate/Turnitin : 17%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.

Palembang, Juli 2022

Muhammad Arun Nugraha

NIM.09011181823003

KATA PENGANTAR

Assalamu'alaikum Wr.Wb.

Puji syukur Alhamdulillah penulis panjatkan kehadirat Allah SWT, dengan karunia dan rahmat-Nya, penulis dapat menyelesaikan penulisan laporan Tugas Akhir yang berjudul **"Implementasi Algoritma *Random Forest* dan Seleksi Fitur *Chi-Square* pada Klasifikasi Serangan *Botnet* di Jaringan *Internet of Things* (IoT) ".**

Pada laporan ini, penulis menjelaskan mengenai pemodelan untuk identifikasi dan klasifikasi serangan botnet terhadap suatu publikasi dengan disertai data-data yang diperoleh penulis saat melakukan penelitian data. Penulis berharap agar laporan ini dapat membantu dan bermanfaat bagi orang banyak.

Pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan laporan Tugas Akhir ini. Oleh karena itu, penulis ingin mengucapkan rasa syukur kepada Allah SWT dan terimakasih kepada yang terhormat :

1. Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan penulisan Proposal Tugas Akhir ini dengan baik dan lancar.

2. Orang tua tercinta yang telah membesarkan saya dengan penuh kasih sayang dan selalu mengajarkan saya dalam berbuat hal yang baik. Terimakasih untuk segala do'a, motivasi dan dukungannya baik moril, materil maupun spritual selama ini.

3. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.

4. Bapak Dr. Ir. H. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya

5. Bapak Deris Stiawan, MT., Ph.D., IPU. selaku Dosen Pembimbing Tugas Akhir yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.

6. Bapak Dr. Ir. Bambang Tutuko, M.T., selaku Pembimbing Akademik Jurusan Sistem Komputer.

8. Mbak Reni selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas.

9. Teman-teman dari Kaktus Kodular Squad yaitu Deny, Jonathan, Wahyu, Shiro, Bayu, dan Dhoni yang sering membantu saya dalam bentuk support pada pembuatan tugas akhir ini.

10. Dan semua pihak yang telah membantu.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis. Akhir kata penulis berharap, semoga proposal tugas akhir ini bermanfaat dan berguna bagi khalayak.

Wassalamu'alaikum Wr. Wb.

Indralaya, Juni 2022
Penulis,

Muhammad Arun Nugraha
NIM. 09011181223003

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DAN SELEKSI FITUR *CHI-SQUARE* PADA KLASIFIKASI SERANGAN *BOTNET* DI JARINGAN *INTERNET OF THINGS* (IoT)

Muhammad Arun Nugraha (09011181823003)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer,
Universitas SriwjayaEmail : arunisarn1304@gmail.com

ABSTRAK

Seiring dengan berkembangnya istilah *Internet of Things* (IoT) di masa sekarang, semakin banyak perangkat keras atau elektronik yang terhubung ke internet. Hal ini memungkinkan banyaknya perangkat – perangkat yang berpotensi terkena serangan *botnet*. *Botnet* merupakan salah satu ancaman yang paling sering ditemukan pada sistem dan keamanaan perangkat atau jaringan IoT di era komputasi berbasis *cloud* di zaman modern. Oleh karena itu, sangat penting untuk memahami anatomi pada *botnet*, mengklasifikasi serangan *botnet*, serta mekanisme apa yang bisa digunakan untuk menghadapi serangan berbasi botnet yang terjadi pada perangkat dan jaringan IoT. *Machine Learning* (ML) telah digunakan dalam penelitian sebagai salah satu solusi yang berpotensi dalam menghadapi ancaman serangan *botnet* pada IoT. *Machine Learning* juga membutuhkan seleksi fitur yang dapat membantu mengurangi jumlah fitur yang ada pada dataset dan memlih fitur yang paling cocok untuk melakukan klasifikasi. Penelitian ini menggunakan *Network Dataset* pada dataset *ToN-IoT* yang dikembangkan oleh laboratorium Cyber Range di University of New South Wales, Canberra. Seleksi fitur *Chi-Square* diterapkan untuk menyeleksi fitur terbaik, dan algoritma *Random Forest* digunakan pada proses klasifikasi. Hasil dari klasifikasi yang menggunakan seleksi fitur *Chi-Square* mampu memperoleh tingkat nilai akurasi, presisi, *recall*, spesifisitas, *F1-Score*, dan nilai *error* dengan baik serta dengan tingkat kesalahan klasifikasi yang relative rendah dibanding hasil klasifikasi tanpa seleksi fitur, dimana hasil terbaik memiliki nilai akurasi sebesar 99.83%, presisi sebesar 99.94%, *Recall* sebesar 99.57%, Spesifisitas sebesar 99.97%, *F1-Score* sebesar 99.75%, dan nilai *error* yang rendah yaitu sebesar 0.17%.

Kata Kunci : *Internet of Things, Botnet Attack Classification, Chi-Square, Random Forest, Machine Learning*

IMPLEMENTATION OF RANDOM FOREST ALGORITHM AND CHI-SQUARE FEATURE SELECTION FOR BOTNET ATTACKS CLASSIFICATION ON INTERNET OF THINGS (IoT) NETWORK

Muhammad Arun Nugraha (09011181823003)
Department of Computer Engineering, Faculty of Computer Science, Sriwijaya University
Email: arunisarn1304@gmail.com

ABSTRACT

Along with the development of the term Internet of Things (IoT) in the present day, more and more hardware or electronics are connected to the internet. This allows many devices to be potentially affected by botnet attacks. Botnets are one of the most common threats to systems and the security of IoT devices or networks in the era of cloud-based computing in modern times. Therefore, it is very important to understand the anatomy of botnets, classify botnet attacks, and what mechanisms can be used to deal with botnet-based attacks that occur on IoT devices and networks. Machine Learning (ML) has been used in research as one of the potential solutions in facing the threat of botnet attacks on IoT. Machine Learning also requires feature selection that can help reduce the number of features present in the dataset and choose the most suitable features for classification. This research uses Network Dataset on ToN-IoT Dataset developed by cyber Range Laboratory at University of New South Wales, Canberra. Chi-Square feature selection is applied to select the best features, and Random Forest algorithm is used in the classification process. The results of the classification using Chi-Square feature selection were able to obtain the level of accuracy, precision, recall, specificity, F1-Score, and error values well and with a relatively low classification error rate compared to classification results without feature selection, where the best results have an accuracy value of 99.83%, precision of 99.94%, Recall of 99.57%, specificity of 99.97%, F1-Score of 99.75%, and a low error value of 0.17%.

Keywords : *Internet of Things, Botnet Attack Classification, Chi-Square, Random Forest, Machine Learnin*

DAFTAR ISI

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR TABEL

Halaman

BAB I
PENDAHULUAN

1.1 Latar Belakang

Seiring dengan banyaknya miniaturiasi perangkat keras secara konstan, dan peningkatan efisiensi daya pada perangkat – perangkat keras tersebut di masa sekarang, telah memungkinkan untuk memasukkan kecerdasan pada perangkat keras atau elektronik yang biasa. Hal tersebut telah memunculkan istilah yang disebut *Internet of Things* (IoT) [1]. Karena istilah IoT ini sedang berkembang di masa sekarang, semakin banyak perangkat keras atau elektronik yang terhubung ke internet. Hal ini memungkinkan banyaknya perangkat – perangkat yang berpotensi terkena serangan *botnet* [2].

Botnet merupakan salah satu ancaman yang paling sering ditemukan pada sistem dan keamanaan perangkat atau jaringan IoT di era komputasi berbasi *cloud* di zaman modern. Karena peningkatan yang sangat besar dalam perangkat yang saling terhubung satu sama lain, dan banyaknya sistem perangkat yang menggunakan IoT, jenis dan pola serangan pada *botnet* selalu berubah. Sebagai contoh, sebanyak 143.000 – 225.000 serangan *botnet* IoT *Mirai* terjadi pada tahun 2018 – 2019. Untuk alasan tersebut, sangat penting untuk memahami anatomi pada *botnet*, mengklasifikasi serangan *botnet*, serta mekanisme apa yang bisa digunakan untuk menghadapi serangan berbasi *botnet* yang terjadi pada perangkat dan jaringan IoT [3]. *Forbes* menyatakan bahwa pada tahun 2019, seraangan pada jaringan IoT terjadi sebanyak 2,9 miliar, 3 kali lebih banyak dari tahun sebelumnya. Tidak hanya itu, *SonicWall* Juga menyatakan bahwa serangan *malware* pada IoT meningkat 215,7% dari 10,3 juta pada tahun 2017 menjadi 32,7 juta pada tahun 2018 [4].

Machine Learning (ML) telah digunakan dalam penelitian sebagai salah satu solusi yang berpotensi dalam menghadapi ancaman serangan *botnet* pada IoT, dikaranakan banyaknya data yang dihasilkan dan tersedia untuk perangkat dan jaringan IoT. *Machine Learning* memungkinkan sistem menjadi dinamis dan fleksibel terhadap input data yang baru, dikarenakan mereka dapat “belajar” tanpa secara eksplisit diberi tahu apa yang harus dilakukan. Oleh karena itu, mereka

memiliki potensi yang signifikan untuk dipakai untuk melakukan identifikasi dan klasifikasi serangan *botnet* pada jaringan IoT [4].

Penelitian yang telah dilakukan oleh peneliti sebelumnya melakukan komparasi beberapa metode *Machine Learning* seperti *Random Forest, K-Nearest Neighbor, Naïve Bayes,* dan *Decision Trees* dalam melakukan klasifikasi botnet. Hasil yang didapat ialah metode *Random Forest* dapat melakukan klasifikasi dengan nilai akurasi yang lebih tinggi dibanding metode – metode lainnya. Metode *Random Forest* memiliki kemampuan menangani beberapa bot, menangani data yang telah dilabel dengan berbagai tipe, mudah dalam melakukan *training*nya, dan komputasi yang efisien [5]. Pada penelitian [6], dilakukan klasifikasi botnet menggunakan metode machine learning, dinama K-NN, *Decision Trees*, dan *Random Forest* memperoleh masing-masing akurasi sebesar 90.25%, 93.15%, dan 95.80% .

Seleksi fitur adalah bagian penting dalam *Machine Learning*. Seleksi fitur dapat membantu mengurangi jumlah fitur yang ada pada dataset dan memlih fitur yang paling cocok untuk melakukan klasifikasi. Selain itu, seleksi fitur juga dapat mengurangi waktu komputasi algoritma ketika melakukan klasifikasi. *Chi-Square* merupakan salah satu seleksi fitur yang memperhitungkan tipe data dan variabel yang diinput, dan dapat menyeleksi jumlah fitur tanpa harus mengurangi nilai akurasi pada klasifikasi [7].

Berdasarkan penjelasan – penjelasan tersebut, maka penulis akan melakukan penelitian untuk melakukan klasifikasi serangan *botnet* pada jaringan *Internet of Things*, dimana penulis akan mengaplikasikan pemodelan *Machine Learning* untuk melakukan klasifikasi serangan *botnet* tersebut. Pada penelitian ini, penulis akan menggunakan metode *Machine Learning Random Forest* dan seleksi fitur *Chi-Square*, oleh karena itu penulis melakukan penelitian ini dengan judul "Implementasi Algoritma *Random Forest* dan Seleksi Fitur *Chi-Square* pada Klasifikasi Serangan *Botnet* di Jaringan *Internet of Things* (IoT)".

1.2 Rumusan Masalah

Berdasarkan penelitian [6], metode *Random Forest* telah menunjukkan hasil akurasi yang lebih baik dibanding metode *machine learning* yang lainnya dalam mengklasifikasi serangan *botnet*, yaitu sebesar 95.80%, akan tetapi, penelitian tersebut tidak menggunakan seleksi fitur dalam melakukan klasifikasi. Penelitian [7] telah membuktikan bahwa *Chi-Square* merupakan salah satu seleksi fitur yang memperhitungkan tipe data dan variabel yang diinput, dan dapat menyeleksi jumlah fitur tanpa harus mengurangi nilai akurasi pada klasifikasi, sehingga dapat mengklasifikasikan serangan *botnet* dengan baik.

Berdasarkan latar belakang yang telah dipaparkan, maka terdapat beberapa masalah yang dapat dirumuskan pada penelitian ini, yaitu:

1. Bagaimana cara melakukan klasifikasi menggunakan algoritma *Random Forest* dan seleksi fitur *Chi-Square* pada serangan *botnet* pada jaringan *Internet of Things*?

2. Apa saja Output yang dihasilkan oleh algoritma *Random Forest* dan seleksi fitur *Chi-Square* ketika melakukan klasifikasi pada serangan *botnet*?

3. Bagaimana perbandingan hasil dari klasifikasi serangan *botnet* dengan algoritma *Random Forest* yang tidak menggunakan seleksi fitur dan yang menggunakan seleksi fitur *Chi-Square*?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian tugas akhir, yaitu :

1. Melakukan klasifikasi serangan *botnet* menggunakan metode *Random Forest* dengan seleksi fitur *Chi-square.*

2. Mendapatkan *Output* yang dihasilkan oleh algoritma *Random Forest* dengan seleksi fitur *Chi-square.*

1.4 Manfaat Penelitian

Adapun manfaat dari penulisan tugas akhir ini, yaitu :

1. Dapat mengimplementasikan algoritma *Random Forest* dan seleksi fitur *Chi-Square* dalam melakukan klasifikasi serangan botnet pada jaringan *Internet of Things*.

2. Dapat menghasilkan Output dari algoritma *Random Forest* dan seleksi fitur *Chi-Square* pada serangan *botnet*.

3. Sebagai bahan bacaan dan referensi bagi orang-orang yang sedang melakukan penelitian tentang serangan botnet pada *Internet of Things*.

1.5 Batasan Masalah

Adapun batasan masalah dari tugas akhir ini, yaitu :

1. Algoritma yang digunakan ialah algoritma *Random Forest*, dengan seleksi fitur yang digunakan ialah *Chi-Square*.

2. Dataset yang digunakan pada penelitian ialah dataset *ToN-IoT* yang dibuat oleh laboratorium *Cyber Range* di UNSW Canberra[8].

3. Penelitian ini hanya sebatas simulasi program dengan bahasa pemrograman *Python*.

4. Output yang dihasilkan dari penelitian ini hanya berupa nilai akurasi yang digunakan sebagai tolak ukur untuk melihat tingkat kecocokan author yang sesuai dengan label yang digunakan.

5. Dalam penelitian ini tidak membahas mengenai bagaimana cara pencegahan *botnet* pada jaringan *Internet of Things*.

1.6 Metodelogi Penelitian

Pada tugas akhir ini menggunakan metodelogi sebagai berikut :

1. Metode Studi Pustaka dan Literature

Pada metode ini mencari dan mengumpulkan referensi yang berupa literature yang terdapat pada buku dan internet mengenai serangan *botnet* pada jaringan *Internet of Things*, algoritma *Random Forest* dan fitur seleksi *Chi-Square* yang nantinya berguna untuk proses pengerjaan tugas akhir ini.

2. Metode Konsultasi

Pada metode ini melakukan konsultasi kepada pihak-pihak yang memiliki pengetahuan serta wawasan yang baik dalam mengatasi permasalahan yang ditemui pada penulisan tugas akhir ini.

3. Metode Pembuatan Model

Pada metode ini membuat suatu perancangan pemodelan dengan menggunakan simulasi.

4. Metode Pengujian

Pada metode ini melakukan pengujian terhadap simulasi yang telah dibuat, apakah simulasi tersebut dapat menghasilkan nilai akurasi yang baik atau tidak.

5. Metode Analisa dan Kesimpulan

Hasil dari pengujian pada tugas akhir ini akan dianalisis baik kelebihannya maupun kekurangannya dan juga menganalisis bagaimana proses peningkatan kualitas suatu gambar.

1.7 Sistematika Penulisan

Berikut ini merupakan sistematika yang digunakan dalam penulisan penelitian tugas akhir agar dapat mendeskripsikan bab – bab penelitian yang terstruktur. Adapun susunan penulisan yang digunakan antara lain:

BAB I PENDAHULUAN

Pada bab ini, akan dijelaskan tentang latar belakang, tujuan, manfaat, rumusan masalah, dan batasan masalah dari topik yang diangkat, yaitu mengenai klasifikasi serangan *botnet* pada jaringan IoT menggunakan algoritma *Random Forest* dan seleksi fitur *Chi-Square*.

BAB II TINJAUAN PUSTAKA

Bab ini akan mengumpulkan berbagai sumber yang akan dijadikan sebagai referensi penelitian. Isi dari bab ini ialah *literature review* ysng berkaitan dengan masalah serangan *botnet* pada jaringan IoT dengan menggunakan algoritma *Random Forest* dan seleksi fitur *Chi-Square*.

BAB III METODOLOGI PENELITIAN

Pada bab ini, akan dijelaskan pembahasan secara bertahap dan secara rinci mengenai langkah – langkah yang digunakan untuk melakukan klasifikasi serangan botnet pada jaringan IoT. Bab ini akan menjelaskan pendekatan seleksi fitur *Chi-Square*, yang akan digabungkan dengan algoritma *Random Forest* untuk melakukan klasifikasi.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini, akan menjelaskan tentang hasil dari pengujian yang telah dilakukan pada tahap sebelumnya, dan data yang diuji akan dianalisa menggunakan berbagai teknik serta akan dilakukan validasi hasil.

BAB V KESIMPULAN DAN SARAN

Bab ini menjelaskan kesimpulan dan saran dari hasil penelitian yang telah dilakukan.

DAFTAR PUSTAKA

[1] N. Koroniotis, N. Moustafa, and E. Sitnikova, "Forensics and Deep Learning Mechanisms for Botnets in Internet of Things: A Survey of Challenges and Solutions," *IEEE Access*, vol. 7, pp. 61764–61785, 2019, doi: 10.1109/ACCESS.2019.2916717.

[2] M. Alshamkhany, W. Alshamkhany, M. Mansour, M. Khan, S. Dhou, and F. Aloul, "Botnet Attack Detection using Machine Learning," *Proc. 2020 14th Int. Conf. Innov. Inf. Technol. IIT 2020*, pp. 203–208, 2020, doi: 10.1109/IIT50501.2020.9299061.

[3] S. N. Thanh, M. Stege, P. I. El-Habr, J. Bang, and N. Dragoni, "Survey on botnets: Incentives, evolution, detection and current trends," *Futur. Internet*, vol. 13, no. 8, pp. 1–43, 2021, doi: 10.3390/fi13080198.

[4] M. N. Injadat, A. Moubayed, and A. Shami, "Detecting Botnet Attacks in IoT Environments: An Optimized Machine Learning Approach," *Proc. Int. Conf. Microelectron. ICM*, vol. 2020-Decem, 2020, doi: 10.1109/ICM50269.2020.9331794.

[5] Irfan, I. M. Wildani, and I. N. Yulita, "Classifying Botnet Attack on Internet of Things Device Using Random Forest," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 248, no. 1, 2019, doi: 10.1088/1755-1315/248/1/012002.

[6] A. Guerra-Manzanares, J. Medina-Galindo, H. Bahsi, and S. Nõmm, "MedBIoT: Generation of an IoT botnet dataset in a medium-sized IoT network," *ICISSP 2020 - Proc. 6th Int. Conf. Inf. Syst. Secur. Priv.*, no. March, pp. 207–218, 2020, doi: 10.5220/0009187802070218.

[7] S. Pokhrel, R. Abbas, and B. Aryal, "IoT Security: Botnet detection in IoT using Machine learning," pp. 1–11, 2021, [Online]. Available: http://arxiv.org/abs/2104.02231.

[8] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and Adna N Anwar, "TON-

IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems," *IEEE Access*, vol. 8, pp. 165130–165150, 2020, doi: 10.1109/ACCESS.2020.3022862.

[9] S. Jagannath, "IoT Botnet Detection using Machine Learning Techniques Cybersecurity," *Natl. Coll. Ireland.*, 2021.

[10] A. R. Gad, A. A. Nashat, and T. M. Barkat, "Intrusion Detection System Using Machine Learning for Vehicular Ad Hoc Networks Based on ToN-IoT Dataset," *IEEE Access*, vol. 9, no. October, pp. 142206–142217, 2021, doi: 10.1109/ACCESS.2021.3120626.

[11] S. Lee, A. Abdullah, N. Jhanjhi, and S. Kok, "Classification of botnet attacks in IoT smart factory using honeypot combined with machine learning," *PeerJ Comput. Sci.*, vol. 7, pp. 1–23, 2021, doi: 10.7717/PEERJ-CS.350.

[12] M. G. Karthik and M. B. M. Krishnan, "Hybrid random forest and synthetic minority over sampling technique for detecting internet of things attacks," *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2021, doi: 10.1007/s12652-021-03082-3.

[13] J. H. Nord, A. Koohang, and J. Paliszkiewicz, "The Internet of Things: Review and theoretical framework," *Expert Syst. Appl.*, vol. 133, pp. 97–108, 2019, doi: 10.1016/j.eswa.2019.05.014.

[14] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "Concept and hardware considerations for product-service system achievement in internet of things," *2019 Int. Conf. Wirel. Technol. Embed. Intell. Syst. WITS 2019*, no. April, pp. 19–22, 2019, doi: 10.1109/WITS.2019.8723755.

[15] S. Almutairi, S. Mahfoudh, S. Almutairi, and J. S. Alowibdi, "Hybrid Botnet Detection Based on Host and Network Analysis," *J. Comput. Networks Commun.*, vol. 2020, 2020, doi: 10.1155/2020/9024726.

[16] J. Kim, M. Shim, S. Hong, Y. Shin, and E. Choi, "Intelligent detection of iot

botnets using machine learning and deep learning," *Appl. Sci.*, vol. 10, no. 19, pp. 1–22, 2020, doi: 10.3390/app10197009.

[17] H. Alzahrani, M. Abulkhair, and E. Alkayal, "A multi-class neural network model for rapid detection of IoT botnet attacks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 688–696, 2020, doi: 10.14569/IJACSA.2020.0110783.

[18] T. M. Booij, I. Chiscop, E. Meeuwissen, N. Moustafa, and F. T. H. D. Hartog, "ToN_IoT: The Role of Heterogeneity and the Need for Standardization of Features and Attack Types in IoT Network Intrusion Data Sets," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 485–496, 2022, doi: 10.1109/JIOT.2021.3085194.

[19] N. Moustafa, "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets," *Sustain. Cities Soc.*, vol. 72, p. 102994, 2021, doi: 10.1016/j.scs.2021.102994.

[20] Y. B. Wah, N. Ibrahim, H. A. Hamid, S. Abdul-Rahman, and S. Fong, "Feature selection methods: Case of filter and wrapper approaches for maximising classification accuracy," *Pertanika J. Sci. Technol.*, vol. 26, no. 1, pp. 329–340, 2018.

[21] A. Rachmat, "Survei Penerapan Model Machine Learning Dalam Bidang Keamanan Informasi," *J. Sist. Cerdas*, vol. 2, no. 1, pp. 47–60, 2019, doi: 10.37396/jsc.v2i1.20.

[22] P. A. A. Resende and A. C. Drummond, "A survey of random forest based methods for intrusion detection systems," *ACM Comput. Surv.*, vol. 51, no. 3, 2018, doi: 10.1145/3178582.

[23] M. W. Ahmad, J. Reynolds, and Y. Rezgui, "Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees," *J. Clean. Prod.*, vol. 203, pp. 810–821, 2018, doi: 10.1016/j.jclepro.2018.08.207.

[24] S. Mcelwee, “Probabilistic Clustering Ensemble Evaluation for Intrusion Detection by for the degree of Doctor of Philosophy in Information Assurance College of Engineering and Computing Nova Southeastern University,” no. August, 2018, doi: 10.13140/RG.2.2.13702.83525.

[25] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, “The impact of class imbalance in classification performance metrics based on the binary confusion matrix,” *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.

[26] H. Xie, X. Xiang, N. Liu, and B. Dong, “Blind Adversarial Training: Balance Accuracy and Robustness,” 2020, [Online]. Available: http://arxiv.org/abs/2004.05914.