PERBANDINGAN METODE SELEKSI FITUR PADA SISTEM KLASIFIKASI *BOTNET* IoT MENGGUNKAN ALGORITMA *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh :

M. TAUFIQ QURAHMAN
09011381722092

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2021

LEMBAR PENGESAHAN

PERBANDINGAN METODE SELEKSI FITUR PADA SISTEM KLASIFIKASI *BOTNET* IoT MENGGUNKAN ALGORITMA *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh :

M. TAUFIQ QURAHMAN
NIM. 09011381722092

Palembang, 29 Juli 2021
Mengetahui,

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

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Jumat

Tanggal : 9 Juli 2021

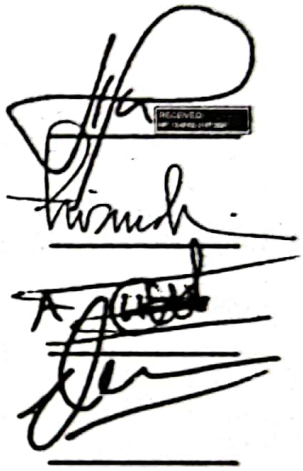
Tim Penguji:

1. Ketua Sidang : Huda Ubaya, M.T

2. Sektretaris Sidang : Tri Wanda Septian M.Sc

3. Penguji Sidang : Ahmad Heryanto, 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 : M. Taufiq Qurahman

NIM : 09011381722092

Judul : Perbandingan Metode Seleksi Fitur Pada Sistem Klasifikasi *Botnet* IoT Menggunakan Algoritma *Random Forest*

Hasil Penyecekan *Software iThenticate/Turnitin* : 4 %

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.

Indralaya, Juli 2021

M. Taufiq Qurahman
NIM. 09011381722092

PERSEMBAHAN

Motto :

“Jadilah Orang Yang Bermanfaat Meskipun Kau Tidak Pintar”
“Jangan Pernah Menyerah Meskipun Kemungkinanya 0,… Sekian Persen Sekalipun”

Ku Persembahkan Untuk :

- ✓ **Kedua Orangtuaku, Kakek dan Nenek, Bibik, dan Seluruh Keluarga Besarku yang tercinta…**
- ✓ **Pihak-pihak yang tidak dapat disebutkan satu persatu dikarenakan akan sangat panjang…..**

Terkhusus untuk Ibuku Erliati Almh yang telah melahirkanku dan merawatku hingga usia enam bulan…
Terkhusus Untuk Nenekku Rohaya Almh dan Kakekku M. Suud yang telah merawatku sedari usia enam bulan…

“Ayah… Ibu… Nenek… Kakek… Bayi yang berusia enam bulan kini telah bersusia 22 tahun dengan menyandang gelar sarjana serta menanggung beban pertanyaan sekarang kerja dimana…???”
“Semoga Ilmu dan Gelar yang ananda peroleh bisa berguna dan menjadi amal jariah untuk kalian… ”

😊

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT, atas segala karunia danrahmat-Nya sehingga penulis dapat menyelesaikan penulisan tugas akhir dengan judul "**Perbandingan Metode Seleksi Fitur Pada Sistem Klasifikasi *Botnet* IoT Menggunakan Algoritma *Random Forest***".

Penulisan laporan tugas akhir ini dilakukan untuk melengkapi salah satu syarat memperoleh gelar Sarjana Komputer di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya. Adapun sebagai bahan penulisan, penulis mengambil berdasarkan hasil penelitian, observasi dan beberapa sumber literatur yang mendukung dalam penulisan laporan ini. Pada kesempatan ini juga, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu baik dari segi moril ataupun materil serta memberikan kemudahan, dorongan, saran dan kritik selama dalam proses penulisan laporan tugas akhir ini.

Oleh karena itu, pada kesempatan ini penulis mengucapkan rasa syukur kepada Allah SWT. dan mengucapkan terima kasih kepada yang terhormat :

1. Orang Tua (Fahrurrozi dan Erliati) serta keluarga besar penulis tercinta, yang telah memberikan doa dan restu serta dukungan yang sangat besar selama mengikuti dan melaksanakan perkuliahan di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya hingga dapat menyelesaikan laporan tugas akhir ini.
2. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Deris Stiawan, Ph. D. selaku Dosen Pembimbing Tugas Akhir penulis sekaligus Dosen Pembimbing Akademik di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya
5. Mbak Nurul Afifah, M.Kom yang telah sangat membantu saya dalam proses pengerjaan laporan tugas akhir ini.

6. Kakak-kakak panutan yang telah memberikan pengetahuan serta saran dan motivasi, Kak Ridho Ilham Renaldo S.kom., Kak Tri Wanda Septian M.Sc
7. Keluarga besar Lab Center of Excellent (CoE) yang juga sangat membantu selama proses pengerjaan tugas akhir
8. Seluruh dosen, staff, serta karyawan Fakultas Ilmu Komputer Universitas Sriwijaya.
9. Abdi Bimantara, Tri Agung Hermawan, Nadhya Hassni, dan Masagus Muhammad Fazri Safiq Riyadhi yang telah membantu Penulis dalam menyelesaikan tugas akhir ini
10. Seluruh teman-teman seperjuangan angkatan 2017 Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
11. Almamater.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan laporan tugas akhir ini. Karena sesungguhnya tak ada yang sempurna didunia ini. Untuk itu, segala saran dan kritik sangatlah penting bagi penulis. Akhir kata, semoga laporan tugas akhir ini dapat bermanfaat dan berguna bagi khalayak.

Indralaya, Juli 2021
Penulis

M. Taufiq Qurahman
Nim. 09011381722092

Perbandingan Metode Seleksi Fitur Pada Sistem Klasifikasi *Botnet* IoT Menggunakan Algoritma *Random Forest*

M. Taufiq Qurahman (09011381722092)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
Email : mtqurahman@gmail.com

Abstrak

Serangan *botnet* menjadi salah satu ancaman yang paling serius dari sekian banyak ancaman dalam pesatnya perkembangan perangkat *Internet of Things* (IoT). Semakin kompleks perangkat IoT membuat waktu proses pendeteksian ataupun pengklasifikasian serangan menjadi lebih lama serta mengkonsumsi banyak memory. Penelitian ini menggunakan dataset *MedBIoT* yang berasal dari Tallinn University Of Technology. Metode seleksi fitur *extra trees* dan *correlation feature selection* diterapkan untuk menyeleksi fitur terbaik. Selain itu, algoritma *random forest* juga diterapkan pada proses klasifikasi. Hasil klasifikasi menggunakan fitur-fitur pilihan mampu memproleh tingkat nilai akurasi, sensitivitas, spesifitas, presisi, dan *F1 score* yang sangat baik dengan waktu proses yang lebih cepat serta dengan tingkat kesalahan klasifikasi yang relatif rendah.

Kata Kunci : *Internet of Things, Botnet Classification, Feature Selection, Random Forest, Machine Learning*

Comparison of Feature Selection Methods in the IoT Botnet Classification System Using Random Forest Algorithm

M. Taufiq Qurahman (09011381722092)
Departement of Computer Engineering, Faculty of Computer Science,
Universitas Sriwijaya
Email : mtqurahman@gmail.com

Abstract

Botnet attacks are one of the most serious threats of many threats in the rapid development of Internet of Things (IoT) devices. The more complex IoT devices make the detection or classifying time of attacks longer and consume a lot of memory. This study used MedBIoT datasets from Tallinn University Of Technology. Extra trees feature selection method and correlation feature selection are applied to select the best features. In addition, the random forest algorithm is also applied to the classification process. Classification results using selected features are able to obtain excellent levels of accuracy, sensitivity, specificity, precision, and F1 scores with faster processing times and with relatively low levels of misclassification.

Keywords : *Internet of Things, Botnet Classification, Feature Selection, Random Forest, Machine Learning*

DAFTAR ISI

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR TABEL

Halaman

BAB I. PENDAHULUAN

Pendahuluan bab ini berisi penjelasan tentang latar belakang penelitian yang berjudul: “Komparasi Metode Seleksi Fitur Pada Sistem Klasifikasi *Botnet* IoT Menggunakan Algoritma *Random Forest*”. *Botnet* IoT merupakan robot jaringan pada kumpulan perangkat yang saling terkoneksi dimana perangakat tersebut telah disusupi program berbahaya ataupun *malware* untuk dikendalikan. IoT menjadi sasaran dari serangan *Botnet* karena IoT merupakan kumpulan dari perangkat-perangkat yang saling terkoneksi melalui jaringan internet [1]. Implementasi metode seleksi fitur digunakan untuk melihat fitur mana saja yang paling berpengaruh terhadap nilai akurasi yang diperoleh dan fitur yang paling berpengaruh tersebut akan diproses sehingga dalam proses klasifikasi menggunakan *Random Forest* akan mengurangi penggunaan memori dan dapat mempercepat waktu proses.

1.1 Latar Belakang

Internet of Things (IoT) merupakan sebuah teknologi komputasi yang mampu menghubungkan setiap perangkat atau objek umum melalui koneksi internet sehingga setiap perangkat maupun objek yang terkoneksi memungkinkan untuk saling bertukar informasi. Salah satu permasalahan keaman yang terjadi pada IoT yaitu ancaman serangan dari *Botnet* [1].

Botnet adalah robot jaringan yang terdapat kumpulan perangkat yang saling terkoneksi melalui internet yang telah disusupi untuk dimata-matai dan dikendalikan dari jarak jauh tanpa sepengetahuan user. *Botnet* umunya terdiri dari seorang *Botmaster* yaitu seorang *cybercriminals* yang bertugas untuk mengontrol, mengoprasikan, dan menjalankan perintah jarak jauh pada *bot. Bot* ialah sebuah perangkat komputer yang telah disusupi yang terhubung dengan perangkat lain melalui protokol apapun [2].

Botnet biasanya menginfeksi seluruh perangkat yang terhubung dengan *malware* seperti virus dan trojan dibawah kendali *botmaster* untuk melakukan pencurian data dan kendali jarak jauh, akan tetapi *botnet* digunakan bukan hanya untuk melakukan pencurian data dan kendali jarang jauh tetapi *botnet* juga digunakan untuk melakukan sabotase serangan dengan sekala yang besar. Menurut penelitian terbaru sebagian besar serangan *Distributed Denial Of Servis* (DDOS) atau serangan penolakan layanan terdistribusi di akibatkan oleh *botnet* dan sebesar 80% hingga 95% serangan *spam* berasar dari *botnet* [3]. Dari berbagai macam permasalahan keamanan yang terjadi pada perangkat *IoT* akibat dari serangan *botnet*, beberapa penelitian mengenai *botnet* pada IoT telah dilakukan.

Seperti pada penelitian [4], mereka menggunakan *fuzzy rule interpolation* untuk mendeteksi *botnet* dan memperoleh hasil akurasi sebesar 95.4%. Algoritma *logistic regression* juga digunakan untuk mengidentifikasi serangan *botnet* dan memperoleh nilai akurasi dari keseluruhan model yang digunakan sebesar 96.7% [5]. Penelitian [6] memproleh presisi 97.7% dan *recall* 99.2% dalam mengklasifikasikan *botnet*. Metode *deep learning* juga di kombinasikan dengan metode seleksi fitur untuk menghasilkan tingkat akurasi tinggi serta pengurangan pemakaian ruang memori sebesar 91,89% [7]. *K*-NN, *Decision Tree*, dan *Random Forest* memperoleh masing-masing akurasi sebesar 90.25%, 93.15%, dan 95.80% dalam klasifikasi *botnet* pada penelitian [8]. Penelitian [9] menganalisa pendeteksian *botnet* menggunakan metode *machine learning* seperti SVM, DT, NB, ANN, dan USML pada dataset UNBS-NB 15 dengan tingkat akurasi yang di peroleh sebesar 84.32%, 94.43%, 71.63%, 63.97%, dan 94.78% . Dari beberapa contoh penelitian yang telah dilakukan terdapat permasalahan pada saat proses klasifikasi dikarenakan ukuran dari data *botnet IoT* yang cukup besar sehingga pada saat memproses penggunaan RAM dan CPU cukup tinggi dan memakan waktu yang cukup lama. Metode *Feature Selection* (Seleksi Fitur) seperti *Correlation Feature Selection* dan *Extra Trees* digunakan untuk memgatasi permasalahan tersebut, diamana dengan menerapkan metode seleksi fitur maka akan diketahui fitur mana saja yang memiliki pengaruh paling tinggi dalam memperoleh hasil akurasi sehingga pada saat proses klasifikasi hanya fitur-fitur yang terbaik yang akan digunakan. Dengan

demikian proses klasifikasi akan sedikit menggunakan RAM dan CPU sehingga akan mempersingkat waktu proses.

Beberapa penelitian yang telah menerapkan metode seleksi fitur seperti [10] telah menggunakan metode seleksi fitur *Correlation Feature Selection* untuk mendeteksi serangan pada perangkat IoT hasilnya sistem pendeteksian yang dikombinasikan dengan metode seleksi fitur tersebut mampu mendeteksi serangan dengan cepat tanpa harus mengorbankan nilai akurasi. *Correlation Feature Selection* juga diterapkan dalam penelitian [11] untuk dikombinasikan dengan metode *machine learning* dalam mendeteksi *botnet* hasilnya waktu komputasi dengan menggunakan *Correlation Feature Selection* jauh lebih cepat dibandingkan waktu komputasi tanpa seleksi fitur yaitu sebesar 8.75 detik dan 17.87 detik menggunakan J48 dan 10.25 detik menggunakan seleksi fitur dan 29.67 detik tanpa seleksi fitur mengunakan *Naïve Bayes*. *Extra Trees* juga digunakan untuk menyeleksi fitur saat mengidentifikasi serangan DDOS agar waktu komputasi meningkat lebih cepat, lima belas fitur terbaik diambil dari jumlah total delapan puluh fitur yang ada dan hasilnya tingkat akurasi sebesar 95% diperoleh pada penelitian [12]. Penelitian [13] menerapkan seleksi fitur *extra trees* untuk menyeleksi fitur terbaik dari total seluruhnya delapan puluh fitur, hasilnya dua puluh delapan fitur dari delapan puluh fitur merupakan fitur terbaik serta akurasi yang diperoleh sebesar 99% dalam waktu 0.29 detik. Penelitian ini dianggap sangat berguna agar dapat mengklasifikasikan *botnet* IoT dengan mengurangi waktu komputasi, pemakain RAM dan CPU, serta meningkatkan akurasi.

1.2 Perumusan Masalah

Botnet menjadi salah satu serangan yang menjadikan perangkat-perangkat IoT sebagai target. Dikarenakan teknologi IoT terdiri dari banyak perangkat yang apabila telah terinfeksi oleh *botnet* maka seluruh perangkat tersebut akan terinfeksi dengan serangan yang sama. Pada saat memproses data serangan *botnet* pada IoT untuk penelitian seperti deteksi [14], identifikasi [15], maupun klasifikasi [16] waktu komputasi untuk penelitian tersebut terbilang cukup lama sehingga memicu pemakaian CPU dan RAM meningkat yang menyebabkan kurang optimalnya pemrosesan.

Berdasarkan latar belang yang telah diuraikan di atas, terdapat beberapa masalah yang dapat dirumuskan dalam penelitian ini yaitu :

1. Bagaimana cara mengoptimalkan waktu komputasi untuk klasifikasi data dengan dimensi yang besar?
2. Bagaimana menerapkan *correlation feature selection* dan *extra trees* untuk menemukan fitur terbaik sehingga mengurangi dimensi data yang besar?
3. Bagaimana pengaruh *correlation feature selection* dan *extra trees* terhadap nilai akurasi, sensitivitas, spesifisitas, presisi, dan *F1 score* dengan kinerja dari sistem klasifikasi *botnet* IoT menggunakan algoritma *Random Forest*?

1.3 Batasan Masalah

Dari perumusan masalah diatas terdapat beberapa batasan masalah dalam penelitian ini yaitu :

1. Penelitian ini hanya menggunakan dataset MedBIoT yang berasal dari *Department of Software Science, Center for Digital Forensics and Cyber Security; Tallinn University of Technology;* Estonia [8].
2. Dataset hanya terdiri dari dua kelas yaitu data *botnet* dan data *legitimate*.
3. Proses seleksi fitur hanya menggunakan *correlation features selection* dan *extra trees.*
4. Peneltian ini hanya berfokus perbandingan hasil antara *correlation feature selection* dan *extra trees* yang kemudian divalidasi menggunakan *random forest*.
5. Penelitian ini tidak membahas cara pencegahan dari serangan tersebut.

1.4 Tujuan

Adapun tujuan dari penelitian adalah sebagai berikut :

1. Menerapkan teknik pemilihan fitur terbaik sehingga dapat menemukan model yang tepat untuk mengoptimalkan waktu komputasi klasifikasi.
2. Menerapkan *correlation feature selection* untuk megetahui fitur terbaik dengan cara merangking subset fitur berbasis korelasi antara fitur dan target dan menerapkan *extra trees* dengan cara menambah pohon keputusan untuk mengetahui rangking dari seluruh fitur berdasarkan dari banyaknya hasil dari pohon keputusan.
3. Mengukur nilai akurasi, presisi, sensitivitas, spesifitas, dan *F1 score* terhadap kinerja *random forest* sebagai pengklasifikasi.

1.5 Manfaat

Adapun hasil dari penelitian ini diharapkan dapat menjadi landasan untuk pengembangan penelitian pada *botnet* IoT melalui metode pemilihan fitur yang tepat agar dapat mengoptimalkan kinerja komputasi. Selain itu beberapa manfaat lain dari penelitian ini antara lain sebagai berikut :

1. Mengetahui teknik pemilihan fitur serta model yang terbaik dalam proses klasifikasi data dengan dimensi yang besar.
2. Mengetahui penerapan teknik pemilihan fitur untuk menyeleksi fitur terbaik dari data dengan dimensi yang besar.
3. Mendapatkan performa terbaik dari hasil seleksi fitur berdasarkan klasifikasi *random forest*.

1.6 Metodologi Penulisan

Agar memperoleh gambaran yang jelas dalam penelitian ini, maka dibuatlah metodologi penulisan yang berisi gambaran dalam tiap bab penelitian ini, yaitu:

1. **BAB I** **Pendahuluan**

   Pada bab ini penulis menjelaskan mengenai latar belakang, perumusan masalah, batasan masalah, tujuan, dan manfaat dari penelitian dengan topik yang dipilih yaitu komparasi metode seleksi fitur pada sistem klasifikasi *botnet* IoT menggunakan algoritma *random forest*.

2. **BAB II** **Tinjauan Pustaka**

   Pada bab ini berisi tentang *literatur review* yang berkaitan dengan masalah *botnet* IoT dengan menggunakan metode seleksi fitur dan algoritma *random forest*.

3. **BAB III** **Metodologi Penelitian**

   Pada bab ini menjelaskan pembahasan secara bertahap dan rinci mengenai langkah yang digunakan untuk mengkomparasi metode seleksi fitur. Metodologi ini menjelaskan pendekatan metode seleksi fitur *boruta* dan *extra trees* serta algoritma *random forest* yang digunakan sehingga tujuan dari penulisan tercapai.

4. **BAB IV** **Analisa dan Pembahasan**

   Pada bab ini berisi tentang hasil dari pengujian yang telah dilakukan serta analisa data yang akan diuji menggunakan berbagai teknik sehingga menemukan model yang tepat dan validasi hasil.

5. **BAB V** **Kesimpulan**

   Pada bab ini berisikan tentang kesimpulan dari hasil penelitian yang telah dilakukan.

DAFTAR PUSTAKA

[1] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, "Industrial internet of things: Challenges, opportunities, and directions," *IEEE Trans. Ind. Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018, doi: 10.1109/TII.2018.2852491.

[2] L. Mathur, M. Raheja, and P. Ahlawat, "Botnet Detection via mining of network traffic flow," *Procedia Comput. Sci.*, vol. 132, pp. 1668–1677, 2018, doi: 10.1016/j.procs.2018.05.137.

[3] C. Yin, Y. Zhu, S. Liu, J. Fei, and H. Zhang, "An enhancing framework for botnet detection using generative adversarial networks," *2018 Int. Conf. Artif. Intell. Big Data, ICAIBD 2018*, pp. 228–234, 2018, doi: 10.1109/ICAIBD.2018.8396200.

[4] M. Al-Kasassbeh, M. Almseidin, K. Alrfou, and S. Kovacs, "Detection of IoT-botnet attacks using fuzzy rule interpolation," *J. Intell. Fuzzy Syst.*, vol. 39, no. 1, pp. 421–431, 2020, doi: 10.3233/JIFS-191432.

[5] R. Bapat *et al.*, "Identifying malicious botnet traffic using logistic regression," *2018 Syst. Inf. Eng. Des. Symp. SIEDS 2018*, pp. 266–271, 2018, doi: 10.1109/SIEDS.2018.8374749.

[6] M. Yusof, M. M. Saudi, and F. Ridzuan, "A new mobile botnet classification based on permission and API calls," *Proc. - 2017 7th Int. Conf. Emerg. Secur. Technol. EST 2017*, pp. 122–127, 2017, doi: 10.1109/EST.2017.8090410.

[7] S. I. Popoola, B. Adebisi, M. Hammoudeh, G. Gui, and H. Gacanin, "Hybrid Deep Learning for Botnet Attack Detection in the Internet of Things

Networks," *IEEE Internet Things J.*, vol. XX, no. X, pp. 1–1, 2020, doi: 10.1109/jiot.2020.3034156.

[8] 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. Icissp, pp. 207–218, 2020, doi: 10.5220/0009187802070218.

[9] T. A. Tuan, H. V. Long, L. H. Son, R. Kumar, I. Priyadarshini, and N. T. K. Son, "Performance evaluation of Botnet DDoS attack detection using machine learning," *Evol. Intell.*, vol. 13, no. 2, pp. 283–294, 2020, doi: 10.1007/s12065-019-00310-w.

[10] M. B. Shahbaz, X. Wang, A. Behnad, and J. Samarabandu, "On efficiency enhancement of the correlation-based feature selection for intrusion detection systems," *7th IEEE Annu. Inf. Technol. Electron. Mob. Commun. Conf. IEEE IEMCON 2016*, 2016, doi: 10.1109/IEMCON.2016.7746286.

[11] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, *Implementing Lightweight IoT-IDS on Raspberry Pi Using Correlation-Based Feature Selection and Its Performance Evaluation*, vol. 926. Springer International Publishing, 2020.

[12] D. V. V. S. Manikumar and B. U. Maheswari, "Blockchain Based DDoS Mitigation Using Machine Learning Techniques," *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020*, pp. 794–800, 2020, doi: 10.1109/ICIRCA48905.2020.9183092.

[13] A. Powell, D. Bates, C. van Wyk, and A. Darren de Abreu, "A cross-comparison of feature selection algorithms on multiple cyber security data-sets," *CEUR Workshop Proc.*, vol. 2540, pp. 196–207, 2019.

[14] F. V. Alejandre, N. C. Cortés, and E. A. Anaya, "Feature selection to detect

botnets using machine learning algorithms," *2017 Int. Conf. Electron. Commun. Comput. CONIELECOMP 2017*, 2017, doi: 10.1109/CONIELECOMP.2017.7891834.

[15] A. Azab, M. Alazab, and M. Aiash, "Machine learning based botnet identification traffic," *Proc. - 15th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 10th IEEE Int. Conf. Big Data Sci. Eng. 14th IEEE Int. Symp. Parallel Distrib. Proce*, pp. 1788–1794, 2016, doi: 10.1109/TrustCom.2016.0275.

[16] H. Bahsi, S. Nomm, and F. B. La Torre, "Dimensionality Reduction for Machine Learning Based IoT Botnet Detection," *2018 15th Int. Conf. Control. Autom. Robot. Vision, ICARCV 2018*, pp. 1857–1862, 2018, doi: 10.1109/ICARCV.2018.8581205.

[17] A. H. Lashkari, G. D. Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features," *ICISSP 2017 - Proc. 3rd Int. Conf. Inf. Syst. Secur. Priv.*, vol. 2017-Janua, no. Cic, pp. 253–262, 2017, doi: 10.5220/0006105602530262.

[18] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018, doi: 10.1016/j.neucom.2017.11.077.

[19] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006, doi: 10.1007/s10994-006-6226-1.

[20] M. A. Hall, "Correlation-based Feature Selection for Machine Learning," no. April, 1999.

[21] 克也嶋崎, "Random Forestsの寄与率を用いた効率的な 特徴選択法の提案," 中部大学工学部情報工学科 卒業論文, pp. 5–32, 2013.

[22] 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.

[23] A. Verikas, E. Vaiciukynas, A. Gelzinis, J. Parker, and M. Charlotte Olsson, "Electromyographic patterns during golf swing: Activation sequence profiling and prediction of shot effectiveness," *Sensors (Switzerland)*, vol. 16, no. 4, 2016, doi: 10.3390/s16040592.

[24] 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.

[25] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, "The balanced accuracy and its posterior distribution," *Proc. - Int. Conf. Pattern Recognit.*, pp. 3121–3124, 2010, doi: 10.1109/ICPR.2010.764.