KLASIFIKASI *TRAFFIC ANONYMOUS THE ONION ROUTE* DENGAN METODE *K-NEAREST NEIGHBOR* (K-NN)

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh :
M. Kabir Akmal
09011382025126

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN

KLASIFIKASI *TRAFFIC ANONYMOUS THE ONION ROUTE* DENGAN *METODE K-NEAREST NEIGHBOR* (K-NN)

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Program Studi Sistem Komputer
Jenjang S1

Oleh:

M. KABIR AKMAL
09011382025126

Palembang, 22 Juli 2024
menyetujui,

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

HALAMAN PERSETUJUAN

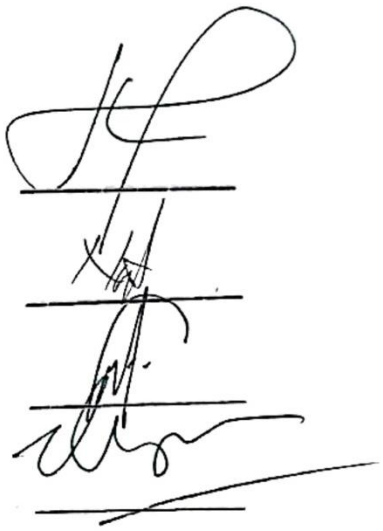
Telah diuji dan lulus pada

Hari : **Senin**

Tanggal : **15 Juli 2024**

Tim Penguji

1. **Ketua** : **Huda Ubaya, M.T.**
2. **Sekretaris** : **Nurul Afifah, M.KOM.**
3. **Penguji** : **Dr. Ahmad Zarkasi, M.T.**
4. **Pembimbing I** : **Prof. Deris Stiawan, M.T., Ph.D.**



Mengetahui, 22/7/24

Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : M. Kabir Akmal

NIM : 09011382025126

Judul : Klasifikasi *Traffic Anonymous The Onion Route* Dengan Metode *K-Nearest Neighbor* (K-NN)

Hasil Pengecekan Plagiat/Turnitin: 10%

Menyatakan bahwa laporann tugas akhir ini adalah hasil karya saya sendiri dan tidak mengandung unsru penjiplakan atau plagiat. Saya sepenuhnya menyadari bahwa jika terbukti adanya penjiplakan atau plagiat dalam skripsi ini, Saya siap menerima sanksi akademik dari Universitas Sriwijaya. Pernyataan ini saya buat dengan kesadaran penuh dan tanpa adanya paksaan dari pihak manapun.

Palembang, Juli 2024

M. Kabir Akmal

NIM. 09011382025126

KATA PENGANTAR

Puji syukur atas kehadirat Allah Subhanahu Wa Ta'ala karena kasih sayang, rahmat, karunia serta bimbingan-Nya, penulis dapat menyelesaikan Skripsi ini. Laporan ini disusun untuk sebagai salah satu mata kuliah yang diberikan kepada mahasiswa Fakultas Ilmu Komputer Program Studi Sistem Komputer.

Adapun skripsi ini berjudul **"KLASIFIKASI *TRAFFIC ANONYMOUS THE ONION ROUTE* DENGAN METODE *K-NEAREST NEIGHBOR* (K-NN)".**

Dalam proses penyusunan Tugas Akhir ini, penulis telah banyak mendapatkan bantuan baik moril maupun materil, bimbingan, sumbangan ide, doa dan lain sebagainya dari berbagai pihak. Untuk itu, pada kesempatan ini penulis ingin menyampaikan dengan segala kerendahan hati, penghargaan yang setinggi-tingginya serta terima kasih sedalam-dalamnya kepada:

1. Allah Subhanahu Wata'ala yang telah memberikan berkah serta nikmat kesehatan dan kesempatan kepada penulis dalam menyusul Tugas Akhir ini.
2. Kedua orang tua penulis yang selalu memberikan doa terbaik serta dukungan secara moril dan materil.
3. Bapak Prof. Dr. Erwin, S.SI, M.SI, selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ir. Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Huda Ubaya, M.T. selaku Sekretaris Jurusan Sistem Komputer Universitas Sriwijaya.
6. Bapak Kemahyanto Exaudi, S.KOM, M.T. selaku Dosen Pembimbing Akademik.
7. Bapak Prof. Deris Stiawan, M.T., Ph.D., IPU., ASEAN Eng selaku Pembimbing Tugas Akhir.
8. Mbak Sari selaku Admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal administrasi selama perkuliahan.
9. Bapak, Ibu dosen jurusan Sistem Komputer yang telah memberikan ilmunya serta pengalamannya kepada saya.
10. Seluruh teman dan sahabat penulis yang selalu memberi dukungan yang tidak bisa penulis sebutkan satu persatu.

11. Teman-teman sekelas SK20 Bukit Universitas Sriwijaya , Terimakasih untuk setiap kebersamaan dan bantuannya selama mengerjakan tugas akhir dan perkuliahan.
12. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan semangat serta do'anya dalam penyelesaian Tugas Akhir.

Kesempurnaan hanya milik Allah dan Rasullnya , Kesalahan dan Kekhilafan pasti selalu ada menghampiri setiap manusia terutama diri saya pribadi. Maka dari itu jikalau dalam penulisan Proposal Tugas akhir ini ini masih terdapat banyak kekurangan dan kesalahan, penulis meminta kritik dan saran yang membangun dengan harapan agar dapat perbaiki di masa yang akan datang, dan semoga tulisan ini dapat bermanfaat bagi semuanya.

Palembang, Juli 2024
Penulis,

M. Kabir Akmal
NIM. 09011382025126

KLASIFIKASI *TRAFFIC ANONYMOUS THE ONION ROUTE* DENGAN METODE *K-NEAREST NEIGHBOR* (K-NN)

M. KABIR AKMAL (09011382025126)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer

Universitas Sriwijaya

Email : kemalll604@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model klasifikasi yang akurat untuk mengidentifikasi dan mengkategorikan lalu lintas anonim pada jaringan Tor. Dataset yang digunakan adalah CIC-Darknet2020. Teknik Synthetic Minority Over-sampling Technique (SMOTE) diterapkan untuk menangani ketidakseimbangan kelas, sementara Mutual Information Classifier (MIC) digunakan untuk seleksi fitur. Model K-Nearest Neighbor (K-NN) digunakan untuk klasifikasi dengan hasil akurasi tertinggi sebesar 93.17% pada n_neighbors = 1. Meskipun tidak menggunakan normalisasi data karena distribusi fitur yang cukup merata, penelitian ini menunjukkan bahwa kombinasi K-NN, SMOTE, dan MIC efektif dalam memberikan klasifikasi lalu lintas anonim yang akurat pada jaringan Tor. Penelitian ini berkontribusi signifikan dalam bidang keamanan jaringan dan dapat menjadi dasar untuk penelitian lebih lanjut dalam deteksi dan keamanan jaringan.

Kata Kunci : Tor, K-NN, CIC-Darknet2020, SMOTE, MIC

CLASSIFICATION TRAFFIC ANONYMOUS THE ONION ROUTE WITH METHOD K-NEAREST NEIGHBOR (K-NN)

M. KABIR AKMAL (09011382025126)

Department of Computer System, Computer Science Faculty

Sriwijaya University

Email : kemall604@gmail.com

ABSTRACT

This research aims to develop an accurate classification model to identify and categorize anonymous traffic on the Tor network. The dataset used is CIC-Darknet2020. The Synthetic Minority Over-sampling Technique (SMOTE) is applied to address class imbalance, while the Mutual Information Classifier (MIC) is employed for feature selection. The K-Nearest Neighbor (K-NN) model is utilized for classification, achieving the highest accuracy of 93.17% with n_neighbors = 1. Despite not using data normalization due to the sufficiently uniform distribution of features, this study demonstrates that the combination of K-NN, SMOTE, and MIC is effective in providing accurate anonymous traffic classification on the Tor network. This research significantly contributes to the field of network security and can serve as a foundation for further research in network detection and security.

Keywords : *Tor, K-NN, CIC-Darknet2020, SMOTE, MIC*

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

Dalam penelitian [1] Anonimitas menjadi salah satu masalah terbesar di dunia keamanan web dan manajemen lalu lintas. Meskipun pengguna web mengkhawatirkan privasi dan keamanan, berbagai metode diterapkan untuk membuat web lebih rentan. Menjelajah web secara anonim tidak hanya mengancam integritas tetapi juga mempertanyakan motivasi aktivitas tersebut.

Dalam penelitian [2] Kemampuan untuk tetap anonim saat terlibat dalam berbagai aktivitas online semakin dicari oleh konsumen yang memiliki masalah privasi. Saat ini satu-satunya cara untuk menjaga anonimitas online adalah melalui penggunaan teknologi. Privasi adalah hak asasi manusia begitu juga dengan privasi online. Meskipun komunikasi dan data memerlukan perlindungan online yang kuat, birokrasi masih lambat dalam merespons laju perubahan teknologi. Kurangnya kepercayaan terhadap domain Teknologi Informasi (TI) telah menyebabkan individu menemukan berbagai cara untuk menyembunyikan identitas online mereka (anonimitas online). Lalu lintas anonim mempersulit pengelolaan dan pemantauan infrastruktur jaringan karena lalu lintas tidak dapat dengan mudah dikaitkan dengan sumber dan atau tujuan aslinya. Dengan kata lain, anonimitas dapat menampung aktivitas kriminal dengan mempersulit penelusuran aktivitas online.

Dalam penelitian [3] Sistem anonimitas jaringan pertama kali diperkenalkan pada awal tahun 1981. Namun, baru kemudian sistem perangkat lunak praktis yang memungkinkan penggunanya berkomunikasi secara anonim di Internet diperkenalkan yaini jaringan anonimisasi Onion route. The Onion Route (ToR) menggunakan skema Perutean Bawang agnostik aplikasi untuk menganonimkan lalu lintas.

Dalam penelitian [4] Tor adalah jaringan anonimasi yang banyak digunakan yang beroperasi berdasarkan protocol Oinon Routing. Tujuan

utamanya adalah untuk menyediakan komunikasi latensi rendah sembari mempertahankan pengguna anonimitas. Dalam penelitian [5] Dengan menggunakan Tor, pengguna dapat mengakses Internet publik tanpa khawatir dengan sensor, pemerintah, penyedia layanan, dan sebagainya.

Dalam penelitian [6] Meluasnya penggunaan browser bawang (Tor) telah menjadi tempat melajunya aktivitas kejahatan siber dan metode identifikasi lalu lintas anonim Tor telah digunakan untuk menyidik lalu lintas web anonim dan mengidentifikasi situs web yang dikunjungi oleh para pelaku kejahatan. Sistem komunikasi anonim menghindari pemantauan dan penyadapan jaringan terhadap informasi pribadi pengguna dengan menyediakan metode teknis untuk menyembunyikan konten dan metadata komunikasi, yang mencakup tiga jenis utama: akses anonim, perutean anonim, dan layanan web gelap. Dalam penelitian [1] Penting untuk mengklasifikasikan lalu lintas jaringan dan mencegah sumber dan tujuan bersembunyi satu sama lain, kecuali untuk aktivitas yang tidak berbahaya.

Pada penelitian [7] Klasifikasi lalu lintas di *Darknet* memainkan peran penting dalam mendeteksi serangan *cyber* dan aktivitas berbahaya di Internet. Ada upaya signifikan telah dilakukan untuk mendeteksi dan mengklasifikasikan lalu lintas terenkripsi dari berbagai darknet dengan mengandalkan teknik pembelajaran mesin. Namun, karena strategi anonimisasi yang digunakan oleh darknet untuk menyembunyikan identitas pengguna, pengenalan lalu lintas praktis merupakan sebuah tantangan. Sistem pengenalan tingkat lanjut didukung dengan teknik kecerdasan buatan untuk memisahkan data lalu lintas *Darknet*.

Pada penelitian [8] Algoritma *K-neighbor* merupakan cara yang efektif untuk mengklasifikasikan sesuatu berdasarkan karakteristiknya, pada penelitian [9] *K-Nearest Neighbour (k-NN)* merupakan suatu metode klasifikasi yang paling dasar dan sederhana yang menjadikan knn harus menjadi salah satu pilihan utama untuk studi klasifikasi ketika pengetahuan tentang distribusi data sedikit atau tidak ada sama sekali. Seringkali berguna untuk menghitung lebih dari satu *Neighbour*, sehingga teknik ini sering disebut dengan Klasifikasi *K-Nearest Neighbour (k-NN)* dimana Klasifikasi *K-Nearest*

Neighbour dapat digunakan untuk menentukan kelasnya. Berdasarkan penelitian [10] ketika estimasi parametrik kepadatan probabilitas yang andal tidak diketahui atau sulit ditentukan. Melalui perkembangan dari kebutuhan Klasifikasi K-Nearest Neighbour digunakan untuk menganalisis diskriminan yang ada.

Dalam penelitian [11] Mengklasifikasikan lalu lintas jaringan penting untuk pembentukan dan pemantauan lalu lintas. Dalam dua dekade terakhir, dengan munculnya permasalahan privasi, pentingnya teknologi yang menjaga privasi semakin meningkat. Jaringan Tor, yang memberikan anonimitas kepada penggunanya dan mendukung layanan anonim yang dikenal sebagai Layanan Onion, adalah cara populer untuk mencapai anonimitas online.

Oleh karena itu, pada tugas akhir ini penulis akan melakukan penelitian yang terfokus pada topik tersebut yaitu dengan judul "*Klasifikasi Traffic Anonymous The Onion Route Dengan Metode K-Nearest Neighbor (K-NN)".*

1.2 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dari penelitian yang berjudul Klasifikasi Traffic Anonymous The Onion Route Dengan Metode K-Nearest Neighbor, antara lain:

1. Melakukan klasifikasi layanan onion untuk lalu lintas darknet.
2. Bagaimana teknik klasifikasi lalu lintas darknet dengan metode K-Nearest Neighbor.
3. Melakukan visualisasi lalu lintas darknet pada klasifikasi layanan onion dengan menggunakan metode K-NN.

1.3 Manfaat

Dari penelitian ini adapun manfaat yang ingin didapat, antara lain :

1. Dapat mengetahui teknik klasifikasi K-Nearest Neighbor pada lalu lintas darknet.

2. Dapat mempelajari kemampuan dan mekanisme klasifikasi layanan onion untuk lalu lintas darknet.

3. Dapat mengidentifikasi lalu lintas anonymous darknet pada

klasifikasi lalu lintas layanan onion dengan menggunakan metode K-Nearest Neighbor.

1.4 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah dalam penelitian tugas akhir ini yaitu, bagaimana Klasifikasi Traffic Anonymous The Onion Route Dengan Metode K-Nearest Neighbour.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini antara lain:

1. Penelitian ini menggunakan dataset CIC-Darknet2020 Internet Traffic.
2. Pada penelitian ini algoritma klasifikasi yang digunakan algoritma K-Nearest Neighbor.
3. Penelitian menggukan *software* jupyter notebook.
4. Visualisasi matriks dilakukan dengan *predicted label* untuk memprediksi setiap kategori atau label dalam data uji.

1.6 Sistematika Penulisan

Sistematika dalam penulisan Tugas Akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bagian bab ini berisi tentang Latar belakang, Tujuan, Manfaat, Rumusan Masalah, Batasan Masalah, dan Sistematika penulisan yang digunakan dalam tugas akhir ini.

BAB II TINJAUAN PUSTAKA

Pada bagian bab ini berisi mengenai bacaan literature yang menjadi referensi serta penjelasan pendukung dari penelitian serta juga landasan teori dari berbagai bahasan yang berhubungan dengan tema.

BAB III METODOLOGI PENELITIAN

Pada bagian bab ini membahas proses penelitian, diagram alir penelitian, serta menjelaskan metodologi penelitian.

BAB IV HASIL DAN ANALISA

Pada bab ini menjelaskan hasil dari penelitian dan analisis *Klasifikasi Traffic Anonymous The Onion Route Dengan Metode K-Nearest Neighbor (K-NN).*

BAB V KESIMPULAN DAN SARAN

Pada bab ini berisi tentang kesimpulan dari penelitian yang dilakukan serta saran untuk penelitian selanjutnya dimasa mendatang.

DAFTAR PUSTAKA

[1] K. M. Akshobhya, "Machine learning for anonymous traffic detection and classification," *Proc. Conflu. 2021 11th Int. Conf. Cloud Comput. Data Sci. Eng.*, no. January, pp. 942–947, 2021, doi: 10.1109/Confluence51648.2021.9377168.

[2] S. Winkler and S. Zeadally, "An analysis of tools for online anonymity," *Int. J. Pervasive Comput. Commun.*, vol. 11, no. 4, pp. 436–453, 2015, doi: 10.1108/IJPCC-08-2015-0030.

[3] S. Chakravarty, A. Stavrou, and A. D. Keromytis, "Identifying proxy nodes in a tor anonymization circuit," *SITIS 2008 - Proc. 4th Int. Conf. Signal Image Technol. Internet Based Syst.*, pp. 633–639, 2008, doi: 10.1109/SITIS.2008.93.

[4] J. A. Vilalonga, J. S. Resende, and H. Domingos, "TorKameleon: Improving Tor's Censorship Resistance With K-anonymization and Media-based Covert Channels," 2023.

[5] O. Galinina, S. Andreev, I. Conference, and D. Hutchison, *Internet of Things , Smart Spaces , and Next Generation*, vol. 1, no. 18. 2018. doi: 10.1007/978-3-030-65729-1.

[6] Y. Lu, M. Cai, C. Zhao, and W. Zhao, "Tor Anonymous Traffic Identification Based on Parallelizing Dilated Convolutional Network," *Appl. Sci.*, vol. 13, no. 5, 2023, doi: 10.3390/app13053243.

[7] M. Alimoradi, M. Zabihimayvan, A. Daliri, R. Sledzik, and R. Sadeghi, "Deep Neural Classification of Darknet Traffic," *Front. Artif. Intell. Appl.*, vol. 356, no. October, pp. 105–114, 2022, doi: 10.3233/FAIA220323.

[8] J. Du, "Global Epidemic Classification Based on K-Nearest Neighbor Algorithm," *Proc. 2021 IEEE Int. Conf. Power Electron. Comput. Appl. ICPECA 2021*, pp. 879–885, 2021, doi: 10.1109/ICPECA51329.2021.9362507.

[9] P. Cunningham and S. J. Delany, "K-Nearest Neighbour Classifiers-A Tutorial," *ACM Comput. Surv.*, vol. 54, no. 6, 2021, doi: 10.1145/3459665.

[10] L. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009, doi: 10.4249/scholarpedia.1883.

[11] I. Karunanayake, N. Ahmed, R. Malaney, R. Islam, and S. K. Jha, “Darknet Traffic Analysis: Investigating the Impact of Modified Tor Traffic on Onion Service Traffic Classification,” *IEEE Access*, vol. 11, pp. 70011–70022, 2023, doi: 10.1109/ACCESS.2023.3293526.

[12] N. Rust-Nguyen, S. Sharma, and M. Stamp, “Darknet traffic classification and adversarial attacks using machine learning,” *Comput. Secur.*, vol. 127, 2023, doi: 10.1016/j.cose.2023.103098.

[13] T. T. T. Nguyen and G. Armitage, “A Survey of Techniques for Internet Traffic Classification using ML,” *Ieee Comst*, vol. 10, no. 4, pp. 56–76, 2008.

[14] M. B. Sarwar, M. K. Hanif, R. Talib, M. Younas, and M. U. Sarwar, “DarkDetect: Darknet Traffic Detection and Categorization Using Modified Convolution-Long Short-Term Memory,” *IEEE Access*, vol. 9, no. Dl, pp. 113705–113713, 2021, doi: 10.1109/ACCESS.2021.3105000.

[15] University of New Brunswick, “CIC-Darknet2020,” 2020. https://www.unb.ca/cic/datasets/darknet2020.html (accessed Jan. 05, 2024).

[16] A. Habibi Lashkari, G. Kaur, and A. Rahali, “DIDarknet: A contemporary approach to detect and characterize the darknet traffic using deep image learning,” *ACM Int. Conf. Proceeding Ser.*, pp. 1–13, 2020, doi: 10.1145/3442520.3442521.

[17] C. Rahalkar, A. Virgaonkar, and K. Varadan, “Analyzing Trends in Tor,” 2022.

[18] A. Khajehpour, F. Zandi, N. Malekghaini, M. Hemmatyar, N. Omidvar, and M. J. Siavoshani, “Deep Inside Tor: Exploring Website Fingerprinting Attacks on Tor Traffic in Realistic Settings,” *2022 12th Int. Conf. Comput. Knowl. Eng. ICCKE 2022*, no. November, pp. 148–156, 2022, doi: 10.1109/ICCKE57176.2022.9960104.

[19] D. L. Huete Trujillo and A. Ruiz-Martínez, “Tor Hidden Services: A Systematic Literature Review,” *J. Cybersecurity Priv.*, vol. 1, no. 3, pp. 496–518, 2021, doi: 10.3390/jcp1030025.

[20] K. Müller, “Defending End-to-End Confirmation Attacks against the Tor Network,” 2015.

[21] A. Ruiz-Martínes, “TOR CIRCUIT,” *ARM’s Privacy Resources.*, 2021. https://webs.um.es/arm/privacy.html (accessed Nov. 12, 2023).

[22] S. H. S. Huang and Z. Cao, “Detecting Malicious Users behind Circuit-Based Anonymity Networks,” *IEEE Access*, vol. 8, pp. 208610–208622, 2020, doi: 10.1109/ACCESS.2020.3038141.

[23] A. Chaabane, P. Manils, and M. A. Kaafar, “Digging into anonymous traffic: A deep analysis of the Tor anonymizing network,” *Proc. - 2010 4th Int. Conf. Netw. Syst. Secur. NSS 2010*, pp. 167–174, 2010, doi: 10.1109/NSS.2010.47.

[24] M. Yang, X. Gu, Z. Ling, C. Yin, and J. Luo, “An active de-anonymizing attack against tor web traffic,” *Tsinghua Sci. Technol.*, vol. 22, no. 6, pp. 702–713, 2017, doi: 10.23919/TST.2017.8195352.

[25] M. Andrus, S. Dean, T. K. Gilbert, N. Lambert, and T. Zick, “AI Development for the Public Interest: From Abstraction Traps to Sociotechnical Risks,” *Int. Symp. Technol. Soc. Proc.*, vol. 2020-Novem, no. Fair ML, pp. 72–79, 2020, doi: 10.1109/ISTAS50296.2020.9462193.

[26] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, *A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science*. 2020. doi: 10.1007/978-3-030-22475-2_1.

[27] A. Shrestha and A. Mahmood, “Review of deep learning algorithms and architectures,” *IEEE Access*, vol. 7, pp. 53040–53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[28] M. P. Hosseini, A. Hosseini, and K. Ahi, “A Review on Machine Learning for EEG Signal Processing in Bioengineering,” *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 204–218, 2021, doi: 10.1109/RBME.2020.2969915.

[29] T. T. T. Nguyen and G. Armitage, “Clustering to assist supervised machine learning for real-time IP traffic classification,” *IEEE Int. Conf. Commun.*, pp. 5857–5862, 2008, doi: 10.1109/ICC.2008.1095.

[30] V. Duarte, S. Zuniga-Jara, and S. Contreras, “Machine Learning and Marketing: A Systematic Literature Review,” *IEEE Access*, vol. 10, no. July, pp. 93273–93288, 2022, doi: 10.1109/ACCESS.2022.3202896.

[31] M. Usama *et al.*, “Unsupervised Machine Learning for Networking:

Techniques, Applications and Research Challenges," *IEEE Access*, vol. 7, pp. 65579–65615, 2019, doi: 10.1109/ACCESS.2019.2916648.

[32] F. Zhao and Q. Tang, "A KNN learning algorithm for collusion-resistant spectrum auction in small cell networks," *IEEE Access*, vol. 6, pp. 45796–45803, 2018, doi: 10.1109/ACCESS.2018.2861840.

[33] Di. C. Nguyen *et al.*, "Enabling AI in Future Wireless Networks: A Data Life Cycle Perspective," *IEEE Commun. Surv. Tutorials*, vol. 23, no. 1, pp. 553–595, 2021, doi: 10.1109/COMST.2020.3024783.

[34] J. Vieira, R. P. Duarte, and H. C. Neto, "Knn-stuff: Knn streaming unit for fpgas," *IEEE Access*, vol. 7, pp. 170864–170877, 2019, doi: 10.1109/ACCESS.2019.2955864.

[35] J. Bascuñana, S. León, M. González-Miquel, E. J. González, and J. Ramírez, "Impact of Jupyter Notebook as a tool to enhance the learning process in chemical engineering modules," *Educ. Chem. Eng.*, vol. 44, no. June, pp. 155–163, 2023, doi: 10.1016/j.ece.2023.06.001.

[36] F. Boukouvala, A. Dowling, J. Verrett, Z. Ulissi, and V. Zavala, "Computational notebooks in chemical engineering curricula," *Chem. Eng. Educ.*, vol. 54, no. 3, pp. 143–150, 2020.

[37] A. Zúñiga-López, C. Avilés-Cruz, A. Ferreyra-Ramírez, and E. Rodríguez-Martínez, "Jupyter-Notebook: A Digital Signal Processing Course Enriched Through the Octave Programming Language," *Adv. Intell. Syst. Comput.*, vol. 1228 AISC, pp. 774–784, 2020, doi: 10.1007/978-3-030-52249-0_52.

[38] J. Wang, T. Y. Kuo, L. Li, and A. Zeller, "Assessing and Restoring Reproducibility of Jupyter Notebooks," *Proc. - 2020 35th IEEE/ACM Int. Conf. Autom. Softw. Eng. ASE 2020*, pp. 138–149, 2020, doi: 10.1145/3324884.3416585.

[39] J. F. Pimentel, L. Murta, V. Braganholo, and J. Freire, "A large-scale study about quality and reproducibility of jupyter notebooks," *IEEE Int. Work. Conf. Min. Softw. Repos.*, vol. 2019-May, pp. 507–517, 2019, doi: 10.1109/MSR.2019.00077.

[40] G. Alfian *et al.*, "Utilizing Random Forest with iForest-Based Outlier Detection and SMOTE to Detect Movement and Direction of RFID Tags,"

Futur. Internet, vol. 15, no. 3, 2023, doi: 10.3390/fi15030103.

[41] A. O. Widodo, B. Setiawan, and R. Indraswari, "Machine Learning-Based Intrusion Detection on Multi-Class Imbalanced Dataset Using SMOTE," *Procedia Comput. Sci.*, vol. 234, pp. 578–583, 2024, doi: 10.1016/j.procs.2024.03.042.

[42] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. June 2002, pp. 321–357, 2002, doi: 10.1613/jair.953.

[43] R. X. W. De and G. Goos, *and Mobile Services – AIMS 2020*. 2020. doi: 10.1007/978-3-030-96033-9.