DISERTASI

DIGITAL FORENSIK PADA JARINGAN IOT SMART HOME DENGAN MACHINE LEARNING

Nama	: ZULHIPNI RENO SAPUTRA ELSI
NIM	: 03013682025017
BKU	: Teknik Informatika
Promotor	: Prof. Ir. Deris Stiawan, S.Kom., M.T., Ph.D
Ko - Promotor	: Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T

PROGRAM STUDI DOKTOR ILMU TEKNIK
FAKULTAS TEKNIK
UNIVERSITAS SRIWIJAYA
2025

DISERTASI

DIGITAL FORENSIK PADA JARINGAN IOT SMART HOME DENGAN MACHINE LEARNING

Nama	**: ZULHIPNI RENO SAPUTRA ELSI**
NIM	**: 03013682025017**
BKU	**: Teknik Informatika**
Promotor	**: Prof. Ir. Deris Stiawan, S.Kom., M.T., Ph.D**
Ko - Promotor	**: Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T**

PROGRAM STUDI DOKTOR ILMU TEKNIK
FAKULTAS TEKNIK
UNIVERSITAS SRIWIJAYA
2025

HALAMAN PENGESAHAN

DISERTASI

(TKT7105)

DIGITAL FORENSIK PADA JARINGAN IOT SMART HOME DENGAN MACHINE LEARNING

Oleh:
ZULHIPNI RENO SAPUTRA ELSI
03013682025017

Telah disetujui
Pada 23 Mei 2025

Promotor

Prof. Ir. Deris Stiawan, S.Kom., M.T., Ph.D. IPU
NIP. 197806172006041002

Ko-Promotor

Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T. IPM
NIP. 197502112003121002

Mengetahui,

Dekan Fakultas Teknik,

Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T. IPM
NIP. 197502112003121002

Koordinator Program Studi

Prof. Dr. Ir. Nukman, M.T.
NIP. 195903211987031001

HALAMAN PERSETUJUAN

Disertasi berjudul "Digital Forensik Pada Jaringan Iot Smart Home Dengan Machine Learning" telah dipresentasikan dihadapan Tim Penguji Disertasi pada Program Studi Doktor Ilmu Teknik Fakultas Teknik Universitas Sriwijaya pada hari Jumat, 23 Mei 2025.

Palembang, 23 Mei 2025
Tim Penguji Disertasi berupa Disertasi:

Ketua Tim Penguji:
Ir. Irsyadi Yani, S.T., M.Eng., Ph.D., IPM ( )
NIP. 197112251997021001

Anggota Tim Penguji:
1. **Prof. Dr. Ir. Nukman, M.T.** ( )
   NIP. 195903211987031001

2. **Prof. Dr. Yusuf Hartono, M.Sc** ( )
   NIP. 196411161990031002

3. **Dr. Kurniabudi, S,Kom., M.Kom** ( )
   NIDN. 1027067601

Mengetahui,

Dekan Fakultas Teknik,

Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T. IPM
NIP. 197502112003121002

Koordinator Program Studi

Prof. Dr. Ir. Nukman, M.T.
NIP. 195903211987031001

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama	: Zulhipni Reno Saputra Elsi
NIM	: 03013682025017
Program Studi	: Doktor Ilmu Teknik
BKU	: Teknik Informatika

Dengan ini menyatakan bahwa disertasi saya dengan judul "Digital Forensik Pada Jaringan IoT Smart Home Dengan Machine Learning", bebas dari fabrikasi, falsifikasi, plagiat, kepengarangan yang tidak sah dan konflik kepentingan dan pengajuan jamak, seperti yang tercantum dalam Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 39 Tahun 2021.

Bilamana ditemukan ketidak sesuaian dengan hal-hal di atas, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan aturan yang berlaku.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Palembang, 23 Mei 2025

menyatakan,

Zulhipni Reno Saputra Elsi

NIM. 03013682025017

DAFTAR RIWAYAT HIDUP

Zulhipni Reno Saputra Elsi, lahir di Palembang pada tanggal 5 November 1980. Anak pertama dari lima bersaudara dari pasangan Alm. Mudjibullah dan Alm. Masnun. Riwayat pendidikan penulis dimulai SD di Muhammadiyah 16 Palembang, kemudian SMP Negeri 7 Palembang, SMA YKPP 1 Plaju, dan Diploma Tiga Teknik Elektronika di Politeknik Negeri Sriwijaya (1999-2002), dilanjutkan dengan Strata Satu Teknik Elektro di Universitas Sriwijaya (2002-2005). Kemudian, penulis menempuh pendidikan Strata Dua di bidang Magister Teknik Informatika di Universitas Bina Darma (2012-2014), dan saat ini sedang menempuh pendidikan Strata Tiga Ilmu Teknik di Universitas Sriwijaya (2020-sekarang). Selain itu, pada tahun 2024, penulis juga menyelesaikan Pendidikan Profesi Insinyur di Institut Teknologi Indonesia.

Dalam perjalanan kariernya, penulis telah menempati berbagai posisi penting baik di dunia pendidikan maupun industri. Saat ini, penulis merupakan Dosen Tetap Program Studi Teknologi Informasi di Universitas Muhammadiyah Palembang sejak 1 Agustus 2019, serta menjabat sebagai Ketua Unit Penjamin Mutu Fakultas Teknik sejak 1 September 2020. Sebelumnya, penulis pernah menjabat sebagai Dosen Tetap Program Studi Teknik Komputer di Amik Sigma (2015–2019), Ketua Prodi Teknik Komputer (Agustus–Oktober 2015), dan Wakil Direktur Bidang Akademik di AMIK Sigma (2015–2019). penulis juga pernah mengabdi sebagai Dosen Tidak Tetap di Universitas Musi Rawas pada Program Studi Teknik Sipil (2011–2015) serta menjabat sebagai Sekretaris Program Studi di institusi yang sama (2012–2015). Selain itu, penulis pernah menjadi Dosen Tetap Prodi Teknik Komputer di STMIK MURA (2010–2015). Pengalaman profesional di industri juga menjadi bagian penting dari karier penulis, yakni sebagai Network Field Operation Engineer di PT. Sampoerna Telekomunikasi Indonesia (2006–2010), serta sebagai Staff BSS-NOM di PT. Indosat, Tbk (2004–2006).

Tahun 2020, penulis memutuskan untuk melanjutkan pendidikan S3 pada Program Studi Doktor Ilmu Teknik Universitas Sriwijaya dengan bidang kajian Teknik Informatika. Bidang penelitian disertasi penulis fokus pada deteksi serangan siber pada jaringan IoT Smart Home menggunakan pendekatan machine learning. Dalam melaksanakan penelitian disertasi, penulis dibimbing dan diarahkan oleh Prof. Ir. Deris Stiawan, S.Kom., MT., Ph.D selaku promotor dan Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T sebagai ko-promotor.

Pada tahun 2023 penulis mendapatkan hiba dari Kemdikbudristek dengan skema penelitian disertasi dengan judul “Digital Forensik Pada Jaringan Iot Smart Home Dengan Machine Learning” sebagai anggota peneliti, dan skema pengbdian kepada Masyarakat dengan judul “Peningkatan Kualitas Kegiatan Belajar Mengajar Menggunakan Teknologi Informasi di Orange Islamic School Palembang Sumatera Selatan” sebagai ketua.

KATA PENGANTAR

Puji syukur kita panjatkan kehadirat Allah SWT, dimana berkat rahmat dan karunia-Nya usulan disertasi ini dapat diselesaikan. Disertasi ini diberi judul "DIGITAL FORENSIK PADA JARINGAN IOT SMART HOME DENGAN MACHINE LEARNING". Disertasi ini merupakan salah satu persyaratan dalam menyelesaikan studi pada Program Studi Doktor Ilmu Teknik Universitas Sriwijaya.

Shalawat serta salam kita panjatkan kepada junjungan besar Nabi Muhammad SAW yang selalu memberikan syafaat dalam perjalanan kita. Meskipun penulis telah berupaya untuk menyajikan disertasi yang terbaik, namun masih terdapat kekurangan, oleh karenanya kritik serta saran yang membangun sangat diperlukan untuk kesempurnaan disertasi ini.

Selesainya disertasi ini tentunya tidak terlepas dari dukungan banyak pihak, oleh karenanya pada kesempatan ini penulis ingin mengucapkan terima-kasih kepada:

1. Keluarga Tercinta Almahum Ayah dan almarhum ibu saya yang selalu memberikan ridho dan do'a restu untuk setiap perjalan dan perjuangan saya selama menempuh pendidikan. Istri dan anak-anakku, atas pengertian dan pengorbanannya, serta adik - adikku yang selalu mendukung.
2. Prof. Ir. Deris Stiawan, S.Kom., MT., Ph.D selaku Promotor yang dengan sabar memberikan banyak sekali dukungan dan pengorbanan yang tidak dapat disebutkan satu persatu.
3. Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T selaku ko-promotor yang telah memberikan dukungan dalam penelitian disertasi ini.
4. Dosen-dosen program Doktor Ilmu Teknik, atas ilmu dan kesabarannya.
5. Teman-teman seperjuangan di Research Group RISTECH (Rumah Inovasi Sains dan Teknologi) yang telah memberikan dukungannya baik moril maupun materil.
6. Teman-teman sejawat Fakultas Teknik Universitas Muhammadiyah Palembang, khususnya Program Studi Teknologi Informasi yang telah memberikan semangat.

Terimakasih atas dukungan yang diberikan baik secara moril dan materil, semoga kebaikannya mendapatkan pahala yang setimpal dari Allah SWT. Akhir kata, semoga usulan disertasi ini telah memenuhi persyaratan dan dapat diterima serta penelitiannya dapat segera dilanjutkan.

Penulis,
Zulhipni Reno Saputra Elsi

RINGKASAN

DIGITAL FORENSIK PADA JARINGAN IOT SMART HOME DENGAN MACHINE LEARNING

Disertasi,

Zulhipni Reno Saputra Elsi; Dipromotori oleh Prof. Ir. Deris Stiawan, S. Kom., MT., Ph.D dan Dr. Ir. Bhakti Yudho Suprapto, S.T., M.T

xix + 204 Halaman, 66 Gambar, 41 Tabel

Disertasi ini mengangkat isu penting mengenai meningkatnya ancaman keamanan siber dalam lingkungan IoT Smart Home (ISH) yang semakin kompleks dan dinamis. Untuk menjawab tantangan tersebut, penelitian ini mengembangkan sebuah model investigasi digital forensik yang memanfaatkan pendekatan machine learning guna mendeteksi dan menganalisis aktivitas serangan di jaringan ISH.

Penelitian ini diawali dengan perumusan masalah terkait sulitnya membedakan aliran data normal dan abnormal dalam jaringan IoT serta keterbatasan dataset dan metode deteksi serangan canggih. Untuk itu, penulis mengembangkan rangkaian metodologi yang terdiri dari beberapa fase utama, yakni: ekstraksi fitur data (Feature Engineering), balancing data, dan seleksi fitur menggunakan berbagai pendekatan algoritmik seperti Chi-Square, Mutual Information, serta kombinasi fungsi logika (AND dan OR).

Dataset yang digunakan dalam penelitian ini adalah MQTT-IoT-IDS2020, yang mencakup lima jenis serangan siber, seperti brute-force attack, UDP scan, dan SSH scanning. Proses klasifikasi dilakukan dengan berbagai algoritma pembelajaran mesin, termasuk Decision Tree, Random Forest, Linear Discriminant Analysis, AdaBoost dan XGBoost, lalu dievaluasi menggunakan metrik akurasi, precision, recall, F1-score, G-Mean dan waktu pelatihan.

Hasil penelitian menunjukkan bahwa kombinasi optimal antara metode seleksi fitur dan algoritma klasifikasi mampu mencapai akurasi hingga 100% pada kondisi tertentu. Salah satu temuan penting adalah bahwa pendekatan Bidirectional Feature dengan balancing RCS dan seleksi fitur Chi-Square mampu meningkatkan efisiensi dan performa klasifikasi secara signifikan.

Secara keseluruhan, disertasi ini berhasil menyajikan model digital forensik yang tidak hanya efektif dalam mendeteksi serangan pada jaringan IoT Smart Home, tetapi juga efisien secara komputasi. Model ini diharapkan dapat menjadi acuan dalam pengembangan sistem keamanan berbasis pembelajaran mesin dan memperkuat proses investigasi digital forensik di masa mendatang.

Kata Kunci: Digital Forensik, Smart Home, Feature Extraction, Teknik Balancing Data, Feature Selection

Kepustakaan: 201 (2011 – 2025)

DAFTAR ISI

Halaman

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1. 1. Latar Belakang

Internet of Things (IoT) terus berkembang sebagai ekosistem yang semakin luas dan kuat, mencakup berbagai perangkat pintar yang saling terhubung [1]. IoT telah menjadi komponen penting dalam berbagai aspek kehidupan, dengan penerapan luas di berbagai sektor seperti industri [2], kesehatan [3], transportasi [4], *smart city* (SC) [5], dan *smart home* (SH) [6]. Penelitian oleh S. Kim, [7] menunjukkan bahwa perangkat seperti kamera dan lampu yang terhubung dalam jaringan IoT dapat menghasilkan data yang relevan untuk keperluan investigasi digital forensik.

Penelitian oleh C. Fu, dkk [8] mengusulkan *prototype Home Automation Watcher* (*HAWatcher*) yang dirancang untuk mendeteksi anomali semantik dalam sistem smart home. Sistem ini diuji menggunakan empat testbed yang berasal dari log peristiwa dan data semantik, menghasilkan akurasi sebesar 97,83%. Sementara itu, studi oleh M. Gajewski, dkk [9] memperkenalkan konsep keamanan SH pada *Home Area Networks* (HAN) dengan pendekatan statistik terhadap lalu lintas pada *Home Gateway* (HG), yang memungkinkan deteksi anomali di jaringan HAN.

Dalam penelitian A. Goudbeek, dkk [10], peneliti mengembangkan kerangka kerja forensik untuk sistem otomatisasi rumah (*Home Automation Systems/*HAS) yang terdiri atas tujuh fase, yaitu: (1) persiapan di luar lokasi, (2) pencarian lokasi HAS, (3) pelestarian HAS, (4) identifikasi spesifikasi HAS, (5) pemeriksaan keamanan, (6) akuisisi bukti, dan (7) analisis bukti digital. Kerangka tersebut diuji

melalui tiga studi kasus untuk menunjukkan efektivitasnya dalam penyelidikan forensik. Sementara itu, F. I. Fagbola dan H. S. Venter [11] mengembangkan kerangka kerja *Shadow Internet of Things Digital Forensic Readiness* (*SIoTDFR*) yang terdiri dari enam fase: koneksi perangkat, identifikasi perangkat, pemantauan perangkat shadow IoT, pengumpulan bukti digital, pelestarian bukti digital, dan penyimpanan bukti digital secara aman.

Penelitian oleh S. Costantini, dkk [12] memanfaatkan metode pembelajaran mesin untuk menganalisis bukti yang diperoleh dari perangkat elektronik. Menurut I. Goni, dkk [13], pembelajaran mesin dapat digunakan sebagai alat dalam sistem keamanan siber, mencakup keamanan jaringan, data, *endpoint*, identitas, *cloud*, *IoT*, dan *fog computing*. Dalam studi oleh P. N. Dawadi, dkk [14], pembelajaran mesin digunakan untuk mengekstraksi fitur dari sensor pada *smart home*. Proses ekstraksi dilakukan dengan algoritma *Principal Component Analysis* (*PCA*), menghasilkan 35 fitur dengan satu fitur sebagai label kelas. Selanjutnya, proses klasifikasi dilakukan dengan *Support Vector Machine* (*SVM*) terhadap 263 sampel aktivitas SH, menghasilkan nilai *ROC* sebesar 0,80 dan G-mean sebesar 0,73 pada dua kelas, yaitu penderita demensia dan individu sehat.

Penelitian I. Cvitic, dkk [15] mengembangkan dataset dari data primer dan sekunder menggunakan 41 perangkat *IoT*. Setelah proses ekstraksi fitur statistik, diperoleh 13 fitur yang digunakan dalam klasifikasi berbasis pembelajaran mesin. Evaluasi menghasilkan akurasi 99,79%, *precision* 0,997–0,999, *F-measure* 0,997–0,999, *True Positive Rate* (*TPR*) 0,997–0,999, *False Positive Rate* (*FPR*) 0–0,001, dan *koefisien kappa* sebesar 0,9973. Peneliti Z. Liouane, dkk [16] menekankan

pentingnya tahap pra-pemrosesan untuk menyaring dan memformat data mentah dari lingkungan cerdas agar dapat digunakan secara optimal. Berdasarkan G. Spanos, dkk [17], pra-pemrosesan dilakukan melalui ekstraksi fitur statistik seperti nilai minimum, maksimum, rata-rata, kuartil, median, *standar deviasi*, dan *interkuartil range.* Sementara itu, N. Bolleddula, dkk [18] memilih fitur seperti waktu minimum/maksimum, durasi, jumlah peristiwa, lokasi dominan/sebelumnya, dan sensor dominan/sebelumnya. Penelitian oleh T. Li, dkk [19] menggunakan T-Shark untuk mengekstraksi 87 fitur dari dataset *PCAP*, sedangkan T.-H. Tan, dkk [20] melakukan pra-pemrosesan pada dataset format *TXT* dengan metode segmentasi berdasarkan label dan konversi ke citra *RGB*.

Untuk meningkatkan performa algoritma klasifikasi, terutama dari segi waktu dan akurasi, digunakan metode *Feature Selection* (FS) seperti *Chi-Square* (Chi), yang memberi peringkat fitur berdasarkan nilai *chi-score* [21]. Menurut Q. He dan M. von Davier, [22], *Chi-Square* efektif dalam menemukan kata kunci penting serta menguji kesamaan antara kumpulan teks. Penelitian V. Gaur dan R. Kumar [23] pada dataset *CICDDoS2019* menghasilkan 86 fitur, dengan beberapa fitur (*SourceIP*, *Source Port*, *DestinationIP*, *Destination Port*, *Protocol, FlowID*, dan *Timestamp*) dikecualikan dalam proses FS. Peneliti I. Jahan Ratul, dkk [24] menggunakan algoritma *Chi* untuk memberi peringkat pada 58 fitur dari hasil pra-pemrosesan, dengan lima fitur teratas yaitu: PLT_recovery = 20390478.33, ANC_recovery = 5996503.15, time_to_acute_GvHD_III_IV = 425033.15, survival_time = 82924.01, dan recipient_body_mass = 115.24.

Dalam buku A. Géron [25] disebutkan bahwa soft voting lebih disarankan dibandingkan hard voting untuk estimasi akhir karena dapat menghasilkan skor probabilitas bagi tiap kelas, yang penting dalam konteks pembelajaran mesin. Studi oleh S. K. Bhoi, dkk [26] menggunakan algoritma *K-Nearest Neighbors* (K-NN) dan *Decision Tree* untuk klasifikasi dataset berjumlah 20.000 contoh. Hasil menunjukkan bahwa kedua algoritma ini cocok digunakan pada dataset berukuran kecil. Penelitian M. Al-Hawawreh, dkk [27] mengevaluasi dataset X-IIoTID dengan berbagai algoritma klasifikasi seperti *Decision Tree, Naive Bayes, K-NN*, *Support Vector Machine*, *Logistic Regression*, *Deep Neural Networks*, dan *Gated Recurrent Units*. Data dibagi dalam tiga kategori: kelas *biner* (normal vs serangan), dan dua jenis *multiclass* (dengan 9 dan 18 jenis serangan). *Naive Bayes* menunjukkan performa terendah karena pendekatannya yang mengasumsikan independensi fitur. Sebaliknya, *Decision Tree* memberikan hasil terbaik pada semua kategori, dengan akurasi mencapai 99,54% pada klasifikasi *biner* dan 99,45% pada klasifikasi dengan 18 jenis serangan.

Dari berbagai penelitian tersebut, dapat disimpulkan bahwa meskipun penerapan machine learning telah terbukti efektif dalam mendukung analisis forensik digital pada jaringan *IoT smart home*, masih terdapat tantangan krusial yang perlu diatasi, khususnya ketika berhadapan dengan data berskala besar. Pertama, pemilihan fitur yang relevan dan efektif sangat penting untuk menghindari kebisingan data yang dapat menurunkan performa klasifikasi. Kedua, data IoT umumnya tidak seimbang (*inbalance*), di mana jumlah data normal jauh lebih besar dibandingkan data *anomali* atau serangan, sehingga dibutuhkan teknik *balancing*

data yang tepat agar model tidak bias. Ketiga, proses seleksi fitur yang efisien perlu dikembangkan untuk mempercepat proses pelatihan tanpa mengorbankan akurasi. Dan keempat, peningkatan performa model, baik dari sisi waktu pemrosesan maupun akurasi klasifikasi, menjadi kunci untuk penerapan nyata sistem digital forensik di lingkungan *smart home*.

1. 2. Perumusan Masalah

Kombinasi forensik digital dan pembelajaran mesin memiliki potensi untuk berhasil yang dapat bermanfaat bagi manusia di era informasi. Pembelajaran mesin telah digunakan sebagian besar dalam forensik berbasis analisis gambar. Namun, pembelajaran mesin untuk analisis data bentuk tekstual.

Jaringan *IoT Smart home (ISH)* terdapat masalah-masalah, untuk mengatasi masalah ini dibutuhkan kerangka kerja sehingga dapat membedakan secara efektif alirannormal dan abnormal. Masalah tersebut dipisahkan menjadi tiga sub-masalah utama berikut: 1) mengumpulkan dataset jaringan *IoT* realistis yang akan digunakan untuk melatih dan mengevaluasi model pembelajaran mesin merupakan tantangan besar [28]. Dataset perlu menangkap perilaku normal perangkat, baik PC maupun perangkat *IoT* yang terhubung ke jaringan saat ini; 1) mendeteksi skenario serangan canggih yang menargetkan sistem *IoT* adalah tantangan besar pertama [29]; 3) menyelidiki dan menganalisis serangan jaringan dengan memanfaatkan Pembelajaran Mesin [30-31]. Setiap model pembelajaran mesin paling cocok untuk menangani masalah dan data klasifikasi tertentu. Model yang dipilih harus mampu memproses data dalam jumlah besar secepat mungkin, tanpa mengorbankan

akurasinya. Selain itu, *hyperparameters* model harus dipilih dengan hati-hati untuk memaksimalkan akurasinya, sambil menghindari *overfitting*.

Berdasarkan sub-masalah di atas dan latar belakang teknis, maka permasalahan dalam penelitian ini yang akan diselesaikan:

1) Bagaimana metode ekstraksi fitur yang paling efektif dapat diidentifikasi dan diimplementasikan untuk meningkatkan kualitas representasi data pada sistem jaringan IoT Smart Home berskala besar?
2) Bagaimana pendekatan balancing data yang tepat dapat diterapkan untuk mengatasi ketidakseimbangan kelas dalam dataset jaringan IoT Smart Home berskala besar guna meningkatkan kinerja model klasifikasi?
3) Bagaimana strategi seleksi fitur yang efisien dapat dirancang untuk mengurangi kompleksitas dimensi data tanpa mengorbankan akurasi pada sistem jaringan IoT Smart Home berskala besar?
4) Bagaimana optimasi teknik pemrosesan data dan pemilihan model dapat meningkatkan performa sistem deteksi pada jaringan IoT Smart Home berskala besar?

1. 3. Tujuan Penelitian

Dalam penelitian ini tujuan umum yaitu untuk merancang deteksi serangan pada jaringan ISH dengan menggunakan ekstraksi fitur dan filter Methods sehingga dapat meminimalkan fitur yang diperlukan untuk klasifikasi dalam mengatasi masalah skalabilitas dan sumber daya komputasi di lingkungan ISH. Kemudian secara khusus tujuan dari penelitian ini yaitu:

1) Merancang fitur ekstraksi yang efektif pada jaringan IoT Smarthome dengan data skala besar
2) Merancang teknik balancing pada jaringan IoT Smarthome dengan data skala besar
3) Merancang Teknik seleksi fitur yang efisien pada jaringan IoT Smarthome dengan data skala besar
4) Merancang teknik klasifikasi machine learning Pada jaringan IoT Smarthome dengan data skala besar

1. 4. Ruang Lingkup penelitian

Berdasarkan tujuan yang ingin dicapai sesuai dengan latar belakang masalah dan perumusan masalah maka diperlukan pembatasan ruang lingkup penelitian, antara lain:

1) Menganalisis dataset ISH dengan mengubah data raw dengan ekstensi PCAP melalui ekstraksi fitur.
2) Menentukan fitur dengan algoritma Chi, MI, AND, dan OR dalam pemodel pembelajaran mesin.
3) Mengembangkan pembelajaran mesin, kerangka kerja investigasi jaringan ISH.
4) Tidak membahas sistem pencegahan serangan Intrusion Preventive System (IPS)

1. 5. Sistematika penulisan

Sistematika penulisan dalam karya ilmiah ini terdiri dari tujuh bab. Bab I Pendahuluan mencakup latar belakang, perumusan masalah, tujuan, ruang lingkup,

dan sistematika penulisan. Bab II Landasan Teori membahas teori terkait, seperti digital forensik, IoT smart home, machine learning (supervised dan unsupervised), serta tahap preprocessing data seperti feature extraction, normalisasi, dan balancing. Bab III Metodologi Penelitian menjelaskan kerangka kerja, sumber data, arsitektur jaringan MQTT, dataset, lingkungan eksperimen, dan metode validasi.

Bab IV membahas hasil dan eksperimen pada metode feature engineering, unidirectional, dan bidirectional. Bab V mengulas balancing data menggunakan SMOTE dan RCS untuk dataset berskala besar. Bab VI menjelaskan proses dan hasil seleksi fitur dengan metode Chi-Squared, Mutual Information, serta kombinasi AND dan OR. Terakhir, Bab VII Penutup berisi simpulan dari hasil penelitian dan saran untuk pengembangan ke depan.

DAFTAR PUSTAKA

[1] I. Idrissi, M. Boukabous, M. Azizi, O. Moussaoui, and H. El Fadili, "Toward a deep learning-based intrusion detection system for IoT against botnet attacks," *IAES Int. J. Artif. Intell.*, vol. 10, no. 1, pp. 110–120, 2021, doi: 10.11591/ijai.v10.i1.pp110-120.

[2] Q. Zhang, L. T. Yang, Z. Chen, P. Li, and F. Bu, "An Adaptive Dropout Deep Computation Model for Industrial IoT Big Data Learning With Crowdsourcing to Cloud Computing," *IEEE Trans. Ind. INFORMATICS*, vol. 15, no. 4, pp. 2330–2337, 2019, doi: 10.1109/TII.2018.2791424.

[3] A. Aldahiri, B. Alrashed, and W. Hussain, "Trends in Using IoT with Machine Learning in Health Prediction System," *forecasting*, vol. 3, no. 1, pp. 181–206, 2021, doi: https://doi.org/10.3390/forecast3010012.

[4] F. Zantalis, G. Koulouras, S. Karabetsos, and D. Kandris, "A Review of Machine Learning and IoT in Smart Transportation," *Futur. internet*, vol. 11, no. 4, pp. 1–23, 2019, doi: https://doi.org/10.3390/fi11040094.

[5] J. Chin, V. Callaghan, and I. Lam, "Understanding and Personalising Smart City Services Using Machine Learning, the Internet-of-Things and Big Data," in *International Symposium on Industrial Electronics*, 2017, pp. 2050–2055, doi: 10.1109/ISIE.2017.8001570.

[6] S. Peter and R. K. Gopal, "Multi-level authentication system for smart home-security analysis and implementation," 2016, doi: 10.1109/INVENTIVE.2016.7824790.

[7] S. Kim, M. Park, S. Lee, and J. Kim, "Smart Home Forensics—Data Analysis of IoT Devices," *Electronics*, vol. 9, no. 8, 2020, doi: 10.3390/electronics908121.

[8] C. Fu, Q. Zeng, and X. Du, "HAWatcher: Semantics-Aware Anomaly Detection for Appified Smart Homes," in *30th USENIX Security Symposium*, 2021, pp. 4223–4240, [Online]. Available: https://www.usenix.org/system/files/sec21-fu-chenglong.pdf.

[9] M. Gajewski, J. M. Batalla, G. Mastorakis, and C. X. Mavromoustakis, "Anomaly traffic detection and correlation in Smart Home automation IoT systems," *Trans. Emerg. Telecommun. Technol.*, p. e4053, 2020, doi: https://doi.org/10.1002/ett.4053.

[10] A. Goudbeek, K.-K. R. Choo, and N.-A. Le-Khac, "A Forensic Investigation Framework for Smart Home Environment," in *7th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering*, 2018, pp. 1446–151, doi: 10.1109/TrustCom/BigDataSE.2018.00201.

[11] F. I. Fagbola and H. S. Venter, "Smart Digital Forensic Readiness Model for Shadow IoT Devices," *Appl. Sci.*, vol. 12, pp. 1–19, 2022, doi: 10.3390/app12020730.

[12] S. Costantini, G. De Gasperis, and R. Olivieri, "Digital forensics and investigations meet artificial intelligence," *Ann. Math. Artif. Intell.*, vol. 86,

no. 1, pp. 193–229, 2019, doi: 10.1007/s10472-019-09632-y.
[13] I. Goni, J. M. Gumpy, T. U. Maigari, M. Muhammad, and A. Saidu, "Cybersecurity and Cyber Forensics: Machine Learning Approach," *Mach. Learn. Res.*, vol. 5, no. 4, pp. 46–50, 2020, doi: 10.11648/j.mlr.20200504.11.
[14] P. N. Dawadi, D. J. Cook, M. Schmitter-Edgecombe, and C. Parsey, "Automated assessment of cognitive health using smart home technologies," *echnology Heal. Care*, vol. 21, no. 4, pp. 323–343, 2013, doi: 10.3233/THC-130734.
[15] I. Cvitic, D. Perakovic, M. Perisa, and B. Gupta, "Ensemble machine learning approach for classifcation of IoT devices in smart home," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 11, pp. 3179–3202, 2021, doi: 10.1007/s13042-020-01241-0.
[16] Z. Liouane, T. Lemlouma, P. Roose, F. Weis, and H. Messaoud, "An improved extreme learning machine model for the prediction of human scenarios in smart homes," *Appl. Intell.*, vol. 48, no. 8, pp. 2017–2030, 2018, doi: 10.1007/s10489-017-1062-5.
[17] G. Spanos, K. M. Giannoutakis, K. Votis, and D. Tzovaras, "Combining Statistical and Machine Learning Techniques in IoT Anomaly Detection for Smart Homes," 2019, doi: 10.1109/CAMAD.2019.8858490.
[18] N. Bolleddula, G. Y. C. Hung, D. Ma, H. Noorian, and D. M. Woodbridge, "Sensor Selection for Activity Classification at Smart Home Environments," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 3927–3930, doi: 10.1109/EMBC44109.2020.9176631.
[19] T. Li, Z. Hong, and L. Yu, "Machine Learning-based Intrusion Detection for IoT Devices in Smart Home," in *2020 IEEE 16th International Conference on Control & Automation (ICCA)*, 2020, pp. 277–282, doi: 10.1109/ICCA51439.2020.9264406.
[20] T.-H. Tan, M. Gochoo, S.-C. Huang, Y.-H. Liu, S.-H. Liu, and Y.-F. Huang, "Multi-Resident Activity Recognition In A Smart Home Using RGB Activity Image and DCNN," *Sensors*, vol. 18, no. 23, pp. 9718–9727, 2018, doi: 10.1109/JSEN.2018.2866806.
[21] T. D. Diwan *et al.*, "Feature Entropy Estimation (FEE) for Malicious IoT Traffic and Detection Using Machine Learning," *Mob. Inf. Syst.*, no. Distributed Secure Computing for Smart Mobile IoT Networks 2021, pp. 1–13, 2021, doi: https://doi.org/10.1155/2021/8091363.
[22] Q. He and M. von Davier, "Identifying Feature Sequences from Process Data in Problem-Solving Items with N-Grams BT - Quantitative Psychology Research," in *Springer Proceedings in Mathematics & Statistics*, 2015, pp. 173–190, doi: 10.1007/978-3-319-19977-1_13.
[23] V. Gaur and R. Kumar, "Analysis of Machine Learning Classifiers for Early Detection of DDoS Attacks on IoT Devices," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1353–1374, 2022, doi: 10.1007/s13369-021-05947-3.
[24] I. Jahan Ratul *et al.*, "Survival Prediction of Children Undergoing Hematopoietic Stem Cell Transplantation Using Different Machine Learning Classifiers by Performing Chi-squared Test and Hyper-parameter

Optimization: A Retrospective Analysis," *arXiv e-prints*, p. arXiv:2201.08987, Jan. 2022, doi: 10.1155/2022/9391136.

[25] A. Géron, "Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems," 2019, doi: 1492032611.

[26] S. K. Bhoi *et al.*, "FireDS-IoT: A Fire Detection System for Smart Home Based on IoT Data Analytics," in *International Conference on Information Technology (ICIT)*, 2018, pp. 161–165, doi: 10.1109/ICIT.2018.0004.

[27] M. Al-Hawawreh, E. Sitnikova, and N. Aboutorab, "X-IIoTID: A Connectivity- and Device-agnostic Intrusion Dataset for Industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3962–3977, 2021, doi: 10.1109/JIOT.2021.3102056.

[28] C. Grajeda, F. Breitinger, and I. Baggili, "Availability of datasets for digital forensics – And what is missing," *Digit. Investig.*, vol. 22, pp. S94–S105, 2017, doi: 10.1016/j.diin.2017.06.004.

[29] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and Other Botnets," *Computer (Long. Beach. Calif).*, vol. 5, no. 7, 2017, doi: 10.1109/MC.2017.201.

[30] R. Walsh, D. Lapsley, W. T. Strayer, and C. Livadas, "Usilng Machine Learning Technliques to Identify Botnet Traffic," in *31st IEEE Conference on Local Computer Networks*, 2006, pp. 967–974, doi: 10.1109/LCN.2006.322210.

[31] R. Doshi, N. Apthorpe, and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," in *IEEE Security and Privacy Workshops (SPW)*, 2018, pp. 29–35, doi: 10.1109/SPW.2018.00013.

[32] M. A. Hassan, G. Samara, and M. A. Fadda, "IoT Forensic Frameworks (DFIF, IoTDOTS, FSAIoT): A Comprehensive Study," *Int. J. Adv. Soft Comput. its Appl.*, vol. 14, no. 1, pp. 71–86, 2022, doi: 10.15849/IJASCA.220328.06.

[33] Ahmed MohanRaj Alenezi, "Digital Forensics in the Age of Smart Environments: A Survey of Recent Advancements and Challenges," *Cryptogr. Secur.*, pp. 1–17, 2023, doi: https://doi.org/10.48550/arXiv.2305.09682.

[34] S. Amiroon and C. Fachkha, "Digital Forensics and Investigations of the Internet of Things: A Short Survey," in *2020 3rd International Conference on Signal Processing and Information Security (ICSPIS)*, 2020, pp. 1–4, doi: 10.1109/ICSPIS51252.2020.9340150.

[35] S. Herodotou and F. Hao, "Spying on the Spy: Security Analysis of Hidden Cameras," in *Network and System Security*, Springer Nature Switzerland, 2023, pp. 345--362.

[36] S. Rani, A. Kataria, V. Sharma, S. Ghosh, and ..., "Threats and corrective measures for IoT security with observance of cybercrime: A survey," *... and Mobile Computing*. hindawi.com, 2021, [Online]. Available: https://www.hindawi.com/journals/wcmc/2021/5579148/.

[37] X. Fan, F. Susan, W. Long, and S. Li, "Security Analysis of Zigbee," pp. 1–18, 2017, [Online]. Available:

https://courses.csail.mit.edu/6.857/2017/project/17.pdf.
[38] W. Wang, F. Cicala, S. R. Hussain, E. Bertino, and N. Li, “Analyzing the attack landscape of Zigbee-enabled IoT systems and reinstating users’ privacy,” in *WiSec ’20: Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2020, pp. 133–143, doi: https://doi.org/10.1145/3395351.3399349.
[39] S. Sathwara, N. Dutta, and E. Pricop, “IoT Forensic A digital investigation framework for IoT systems,” in *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2018, p. 10.1109/ECAI.2018.8679017, doi: 10.1109/ECAI.2018.8679017.
[40] M. Chernyshev, S. Zeadally, Z. Baig, and ..., “Internet of things forensics: The need, process models, and open issues,” *IT Prof.*, 2018, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8378977/.
[41] K. Janjua, M. A. Shah, A. Almogren, H. A. Khattak, C. Maple, and I. U. Din, “Proactive Forensics in IoT: Privacy-Aware Log-Preservation Architecture in Fog-Enabled-Cloud Using Holochain and Containerization Technologies,” *Electronics*, vol. 9, no. 7, p. 1172, 2020, doi: 10.3390/electronics9071172.
[42] M. Abuhamad *et al.*, “DL-FHMC: Deep Learning-Based Fine-Grained Hierarchical Learning Approach for Robust Malware Classification,” *IEEE Trans. Dependable Secur. Comput.*, vol. 9, no. 5, pp. 3432–3447, 2022, doi: 10.1109/TDSC.2021.3097296.
[43] S. Qureshi, S. Tunio, F. Akhtar, Ahsan Wajahat, A. Nazir, and F. Ullah, “Network Forensics: A Comprehensive Review of Tools and Techniques,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, 2021, doi: 10.14569/IJACSA.2021.01205103.
[44] V. Ndatinya, Z. Xiao, V. R. Manepalli, K. Meng, and Y. Xiao, “Network forensics analysis using Wireshark,” *Int. J. Secur. Networks*, vol. 10, no. 2, pp. 91–106, 2015, doi: 10.1504/IJSN.2015.070421.
[45] M. Fagan, M. Yang, A. Tan, L. Randolph, and K. Scarfone, “Security Review of Consumer Home Internet of Things (IoT) Products,” *NIST IR 8267 (Initial Public Draft.*, 2019, doi: https://doi.org/10.6028/NIST.IR.8267-draft.
[46] J. Dahmen, B. L. Thomas, D. J. Cook, and X. Wang, “Activity Learning as a Foundation for Security Monitoring in Smart Homes,” *Sensors*, vol. 17, no. 4, p. 737, 2017, doi: https://doi.org/10.3390/s17040737.
[47] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, “A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures,” *IEEE Access*, vol. 7, pp. 82721–82743, 2019, doi: 10.1109/ACCESS.2019.2924045.
[48] S. Perumal, N. M. Norwawi, and V. Raman, “Internet of Things(IoT) digital forensic investigation model: Top-down forensic approach methodology,” in *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, 2015, pp. 19–23, doi: 10.1109/ICDIPC.2015.7323000.
[49] C. Meffert, D. Clark, I. Baggili, and F. Breitinger, “Forensic State Acquisition from Internet of Things (FSAIoT): A General Framework and

Practical Approach for IoT Forensics through IoT Device State Acquisition," 2017, doi: 10.1145/3098954.3104053.
[50] M. E. Alex and R. Kishore, "Forensics framework for cloud computing," *Comput. Electr. Eng.*, vol. 60, pp. 193–205, 2017, doi: https://doi.org/10.1016/j.compeleceng.2017.02.006.
[51] P. H. Rughani, "Artificial Intelligence Based Digital Forensics Framework," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 8, pp. 10–14, 2017, doi: https://doi.org/10.26483/ijarcs.v8i8.4571.
[52] S. Venugopal, G. W. Sathianesan, and R. Rengaswamy, "Cyber forensic framework for big data analytics using Sunflower Jaya optimization-based Deep stacked autoencoder," *Int J Numer Model*, vol. 34, no. 5, pp. 1–18, 2021, doi: https://doi.org/10.1002/jnm.2892.
[53] A. Razaque, M. Aloqaily, M. Almiani, Y. Jararweh, and G. Srivastava, "Efficient and reliable forensics using intelligent edge computing," *Futur. Gener. Comput. Syst.*, vol. 118, pp. 230–239, 2021, doi: 10.1016/j.future.2021.01.012.
[54] M. S. Mazhar *et al.*, "Forensic Analysis on Internet of Things (IoT) Device Using Machine-to-Machine (M2M) Framework," *Electronics*, vol. 11, no. 7, pp. 1–23, 2022, doi: https://doi.org/10.3390/electronics11071126.
[55] L. Babun, A. K. Sikder, A. Acar, and A. S. Uluagac, "IoTDots: A Digital Forensics Framework for Smart Environments," *Comput. Sci.*, 2018, doi: 10.48550/arXiv.1809.00745.
[56] A. Awasthi, H. O. L. Read, K. Xynos, and I. Sutherland, "Welcome pwn: Almond smart home hub forensics," *Digit. Investig.*, vol. 26, p. S38 S46, 2018, doi: 10.1016/j.diin.2018.04.014.
[57] O. Takwa, C. R. Belgacem, and D. Adel, "A New Digital Investigation Frameworks Comparison Method," *Int. J. Comput. Tech.*, vol. 3, no. 4, pp. 6–10, 2016, doi: 10.48550/arXiv.1711.02824.
[58] M. R. Hidayat and I. Riadi, "Investigation of Botnet Attacks using Network Forensic Development Life Cycle Method," *Int. J. Comput. Appl.*, vol. 183, no. 25, pp. 30–36, 2021, doi: 10.5120/ijca2021921632.
[59] E. E.-D. Hemdan and D. . Manjaiah, "An efficient digital forensic model for cybercrimes investigation in cloud computing," *Multimed Tools Appl*, vol. 80, pp. 14255–14282, 2021, doi: 10.1007/s11042-020-10358-x.
[60] G. Kumar, R. Saha, C. Lal, and M. Conti, "Internet-of-Forensic (IoF): A blockchain based digital forensics framework for IoT applications," *Futur. Gener. Comput. Syst.*, vol. 120, pp. 13–25, 2021, doi: 10.1016/j.future.2021.02.016.
[61] S. Khan, A. Gani, A. W. A. Wahab, M. Shiraz, and I. A. Khan, "Network forensics: Review, taxonomy, and open challenges," *J. Netw. Comput. Appl.*, vol. 66, pp. 214–235, 2016, doi: 10.1016/j.jnca.2016.03.005.
[62] R. Hunt and S. Zeadally, "Network Forensics: An Analysis of Techniques, Tools, and Trends," *Computer (Long. Beach. Calif).*, vol. 45, no. 12, pp. 36–43, 2012, doi: 10.1109/MC.2012.252.
[63] N. Moustafa and J. Slay, "RCNF: Real-time Collaborative Network Forensic Scheme for Evidence Analysis," *arXiv*, vol. abs/1711.0, 2017.

[64] N. Koroniotis, N. Moustafa, E. Sitnikova, and J. Slay, “Towards Developing Network forensic mechanism for Botnet Activities in the IoT based on Machine Learning Techniques,” in *International Conference on Mobile Networks and Management*, 2017, pp. 30–44, doi: 10.1007/978-3-319-90775-8_3.

[65] N. Moustafa, J. Slay, and G. Creech, “Novel Geometric Area Analysis Technique for Anomaly Detection Using Trapezoidal Area Estimation on Large-Scale Networks,” *IEEE Trans. Big Data*, vol. 5, no. 4, pp. 481–494, 2017, doi: 10.1109/TBDATA.2017.2715166.

[66] M. Keshk, N. Moustafa, E. Sitnikova, and G. Creech, “Privacy Preservation Intrusion Detection Technique for SCADA Systems,” *arXiv Prepr.*, vol. arXiv:1711, 2017, doi: 10.48550/arXiv.1711.02828.

[67] K. Ruan, J. Carthy, T. Kechadi, and I. Baggili, “Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results,” *Digit. Investig.*, vol. 10, pp. 34–43, 2013, doi: 10.1016/j.diin.2013.02.004.

[68] P. Purnaye and V. Kulkarni, “A Comprehensive Study of Cloud Forensics,” *Arch. Comput. Methods Eng.*, vol. 29, pp. 33–46, 2022, doi: 10.1007/s11831-021-09575-w.

[69] V. R. Kebande and I. Ray, “A Generic Digital Forensic Investigation Framework for Internet of Things(IoT),” in *IEEE 4th International Conference on Future Internet of Things and Cloud*, 2016, pp. 356–362, doi: 10.1109/FiCloud.2016.57.

[70] S. Li, K.-K. R. Choo, Q. Sun, W. J. Buchanan, and J. Cao, “IoT Forensics: Amazon Echo as a Use Case,” *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6487–6497, 2019, doi: 10.1109/JIOT.2019.2906946.

[71] W. Ali, G. Dustgeer, M. Awais, and M. A. Shah, “IoT based Smart Home Security Challenges, Security Requirements and Solutio,” in *23rd International Conference on Automation and Computing (ICAC)*, 2017, pp. 1–6, doi: 10.23919/IConAC.2017.8082057.

[72] B.-C. Chifor, I. Bica, V.-V. Patriciu, and F. Pop, “A security authorization scheme for smart home Internet of Things devices,” *Futur. Gener. Comput. Syst.*, vol. 86, no. C, pp. 740–749, 2018, doi: https://doi.org/10.1016/j.future.2017.05.048.

[73] S. Singh, I.-H. Ra, W. Meng, M. Kaur, and G. H. Cho, “SH-BlockCC: A secure and efficient Internet of things smart home architecture based on cloud computing and blockchain technology,” *Int. J. Distrib. Sens. Networks*, vol. 15, no. 4, pp. 1–18, 2019, doi: 10.1177/1550147719844159.

[74] A. Iqbal *et al.*, “Interoperable Internet-of-Things Platform for Smart Home System using Web-of-Objects and Cloud,” *Sustain. Cities Soc.*, vol. 38, pp. 36–646, 2018, doi: https://doi.org/10.1016/j.scs.2018.01.044.

[75] S. Singh, P. K. Sharma, and J. H. Park, “SH-SecNet: An Enhanced Secure Network Architecture for the Diagnosis of Security Threats in a Smart Home,” *Sustainability*, vol. 9, no. 4, pp. 1–19, 2017, doi: https://doi.org/10.3390/su9040513.

[76] W. Abbass, Z. Bakraouy, A. Baina, and M. Bellafkih, “Assessing the Internet

of Things Security Risks," *J. Commun.*, vol. 14, no. 10, pp. 958–964, 2019, doi: 10.12720/jcm.14.10.958-964.

[77] M. Fakroon, M. Alshahrani, F. Gebali, and I. Traore, "Secure remote anonymous user authentication scheme for smart home environment," *Internet of Things*, vol. 9, pp. 1–20, 2020, doi: https://doi.org/10.1016/j.iot.2020.100158.

[78] R. Paudel, W. Eberle, and L. B. Holder, "Anomaly Detection of Elderly Patient Activities in Smart Homes using a Graph-Based Approach," in *ICDATA'18*, 2018, pp. 163–169, doi: 1-60132-481-2.

[79] D. Zheng, Z. Hong, N. Wang, and P. Chen, "An Improved LDA-Based ELM Classification for Intrusion Detection Algorithm in IoT Application," *Sensors*, vol. 20, pp. 2–19, 2020, doi: 10.3390/s20061706.

[80] F. K. Santoso and N. C. H. Vun, "Securing IoT for Smart Home System," in *IEEE International Symposium on Consumer Electronics (ISCE*, 2015, p. 1+2, doi: 10.1109/ISCE.2015.7177843.

[81] M. A. Hoque and C. Davidson, "Design and Implementation of an IoT-Based Smart Home Security System," *Int. J. Networked Distrib. Comput.*, vol. 7, no. 2, pp. 85–92, 2019, doi: 10.2991/ijndc.k.190326.004; ISSN 2211-7946.

[82] S. Tanwar, P. Patel, K. Patel, S. Tyagi, N. Kumar, and M. S. Obaidat, "An Advanced Internet of Thing based Security Alert System for Smart Home," in *2017 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2017, pp. 25–29, doi: 10.1109/CITS.2017.8035326.

[83] S. Arif, M. A. Khan, S. U. Rehman, M. A. Kabir, and M. Imran, "Investigating Smart Home Security: Is Blockchain the Answer?," *IEEE Access*, vol. 8, pp. 117802–117816, 2020, doi: 10.1109/ACCESS.2020.3004662.

[84] S. Wadhwani, U. Singh, P. Singh, and S. Dwivedi, "Smart Home Automation and Security System using Arduino and IOT," *Int. Res. J. Eng. Technol.*, vol. 5, no. 2, pp. 1357–1359, 2018, doi: 2395-0056.

[85] J. Chhabra and P. Gupta, "IoT based Smart Home design using power and security management," in *2016 1st International Conference on Innovation and Challenges in Cyber Security (ICICCS 2016)*, 2016, pp. 6–10, doi: 10.1109/ICICCS.2016.7542317.

[86] Z. Wang, "Smart home system design based on Internet of things," *Appl. Mech. Mater.*, vol. 602–605, pp. 3808–3812, 2014, doi: 10.4028/www.scientific.net/AMM.602-605.3808.

[87] Y. Jie, J. Y. Pei, L. Jun, G. Yun, and X. Wei, "Smart Home System based on IOT Technologies," in *2013 International Conference on Computational and Information Sciences*, 2013, pp. 1789–1791, doi: 10.1109/ICCIS.2013.468.

[88] P. Xiang, "Design of Smart Home System Based on the Technology of Internet of Things," *Res. J. Appl. Sci. Eng. Technol.*, vol. 4, no. 14, pp. 2236–2240, 2012, doi: 2040-7467.

[89] K. Bing, L. Fu, Y. Zhuo, and L. Yanlei, "Design of an Internet of Things-based Smart Home System," in *2011 2nd International Conference on*

Intelligent Control and Information Processing, 2011, pp. 921–924, doi: 10.1109/ICICIP.2011.6008384.

[90] S. Sachdeva. and A. Ali, "Machine learning with digital forensics for attack classification in cloud network environment," *Int. J. Syst. Assur. Eng. Manag.*, vol. 13, pp. 156–165, 2022, doi: https://doi.org/10.1007/s13198-021-01323-4.

[91] A. Dey, "Machine Learning Algorithms: A Review," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 3, pp. 1174–1179, 2016, doi: 0975-9646.

[92] H.-T. Hsu, G.-J. Jong, J.-H. Chen, and C.-G. Jhe, "Improve IoT Security System of Smart-Home by Using Support Vector Machine," in *2019 IEEE 4th International Conference on Computer and Communication Systems*, 2019, pp. 674–677, doi: 10.1109/CCOMS.2019.8821678.

[93] H. Kang, D. H. Ahn, G. M. Lee, J. Do Yoo, K. H. Park, and H. K. Kim, "IoT network intrusion dataset," 2019, doi: https://dx.doi.org/10.21227/q70p-q449.

[94] D. Caputo, A. Ranieri, L. Verderame, A. Merlo, and L. Caviglione, "Google Home Pcap," 2021, doi: https://dx.doi.org/10.21227/pr94-zk95.

[95] G. Mohi-ud-din, "NSL-KDD," 2018, doi: https://dx.doi.org/10.21227/425a-3e55.

[96] F. Marturana and S. Tacconi, "A Machine Learning-based Triage methodology for automated categorization of digital media," *Digit. Investig.*, vol. 10, pp. 193–204, 2013, doi: 10.1016/j.diin.2013.01.001.

[97] K. S. Hoon, K. C. Yeo, S. Azam, B. Shanmugam, and F. De Boer, "Critical review of machine learning approaches to apply big data analytics in DDoS forensics," in *2018 International Conference on Computer Communication and Informatics*, 2018, pp. 1–5, doi: 10.1109/ICCCI.2018.8441286.

[98] A. MacDermott, T. Baker, and Q. Shi, "Iot Forensics: Challenges for the Ioa Era," in *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, 2018, pp. 1–5, doi: 10.1109/NTMS.2018.8328748.

[99] G. Mokhtari, S. Aminikhanghahi, Q. Zhang, and D. J. Cook, "Fall detection in smart home environments using UWB sensors and unsupervised change detection," *J. Reliab. Intell. Environ.*, vol. 4, pp. 131–139, 2018, doi: 10.1007/s40860-018-0065-2.

[100] E. Anthi, L. Williams, M. Słowinska, G. Theodorakopoulos, and P. Burnap, "A Supervised Intrusion Detection System for Smart Home IoT Devices," *Internet Things J.*, vol. 6, no. 5, pp. 9042–9053, 2019, doi: 10.1109/JIOT.2019.2926365.

[101] J. Hurwitz and D. Kirsch, "Machine Learning For Dummies," 2018, doi: 978-1-119-45495-3.

[102] S. Salcedo-Sanz *et al.*, "Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: a review," *Theor. Appl. Climatol.*, vol. 155, pp. 1–44, 2024, doi: https://doi.org/10.1007/s00704-023-04571-5.

[103] P. Bühlmann and T. Hothorn, "Boosting Algorithms: Regularization, Prediction and Model Fitting," *Stat. Sci.*, vol. 22, no. 4, pp. 477–505, 2007,

doi: 10.1214/07-STS242.
[104] R. E. Schapire, “Explaining AdaBoost,” in *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 37–52.
[105] T. Chengsheng, L. Huacheng, and X. Bing, “AdaBoost typical Algorithm and its application research,” in *2017 3rd International Conference on Mechanical, Electronic and Information Technology Engineering (ICMITE 2017)*, 2017, p. 00222, doi: https://doi.org/10.1051/matecconf/201713900222.
[106] A. Criminisi, J. Shotton, and E. Konukoglu, *Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*. Now Foundations and Trends, 2012.
[107] Raul Rojas, “AdaBoost and the Super Bowl of Classifiers A Tutorial Introduction to Adaptive Boosting,” *TechRxiv*, vol. 1, pp. 1–6, 2024, doi: 10.36227/techrxiv.172107276.63524590/v1.
[108] Z. He, D. Lin, T. Lau, and M. Wu, “Gradient Boosting Machine: A Survey,” *arXiv*, vol. stat.ML, pp. 1–9, 2019, doi: https://doi.org/10.48550/arXiv.1908.06951.
[109] A. Anghel, N. Papandreou, T. Parnell, A. De Palma, and H. Pozidis, “Benchmarking and Optimization of Gradient Boosting Decision Tree Algorithms,” *arXiv*, pp. 1–7, 2019, doi: https://doi.org/10.48550/arXiv.1809.04559.
[110] Rory Mitchell, A. Adinets, T. Rao, and E. Frank, “XGBoost: Scalable GPU Accelerated Learning,” *arXiv*, pp. 1–5, 2018, doi: https://doi.org/10.48550/arXiv.1806.11248.
[111] dmlc XGBoost, “XGBoost Documentation,” *XGBoost Documentation*, 2022. https://xgboost.readthedocs.io/en/stable/ (accessed Sep. 29, 2024).
[112] Z. Ma, J. Guo, S. Mao, and T. Gu, “An interpretability research of the Xgboost algorithm in remaining useful life prediction,” in *2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, 2020, pp. 433–438, doi: 10.1109/ICBASE51474.2020.00098.
[113] B. Zhang, F. K. A. Salem, M. J. Hayes, and T. Tadesse, “Quantitative Assessment of Drought Impacts Using XGBoost based on the Drought Impact Reporter,” *arXiv*, pp. 1–5, 2020, doi: https://doi.org/10.48550/arXiv.2211.02768.
[114] S. Mukhopadhyay, “Advanced Data Analytics Using Python With Machine Learning, Deep Learning and NLP Examples,” 2018, doi: 10.1007/978-1-4842-3450-1.
[115] N. Dogru and A. Subasi, “Traffic Accident Detection Using Random Forest Classifier,” in *2018 15th Learning and Technology Conference (L&T)*, 2018, pp. 40–45, doi: 10.1109/LT.2018.8368509.
[116] B. M. H. A. Allen, A. I. Daood, and W. H, “Poster Abstract: Comparison of Classifiers for Prediction of Human Actions in a Smart Home,” in *2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*, 2018, pp. 287–288, doi:

10.1109/IoTDI.2018.00043.
[117] T. OConnor, R. Mohamed, M. Miettinen, W. Enck, B. Reaves, and A.-R. Sadeghi, "HomeSnitch: Behavior Transparency and Control for Smart Home IoT Devices," in *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, pp. 128–138, doi: 10.1145/3317549.3323409.
[118] F. Alghayadh and D. Debnath, "A Hybrid Intrusion Detection System for Smart Home Security," in *2020 IEEE International Conference on Electro Information Technology (EIT)*, 2020, pp. 319–323, doi: 10.1109/EIT48999.2020.9208296.
[119] H. Tyagi and R. Kumar, "Attack and Anomaly Detection in IoT Networks Using Supervised Machine Learning Approaches," *Rev. d'Intelligence Artif.*, vol. 35, no. 1, pp. 11–21, 2021, doi: 10.18280/ria.350102.
[120] T. Nugroho, M. Nasrun, and C. Setianingsih, "Smart Lamp Control Based on User Behavior For Two Lamps Using K-Nearest Neighbour," in *2019 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, 2019, pp. 123–128, doi: 10.1109/ICAMIMIA47173.2019.9223423.
[121] V. Jakkula and D. J. Cook, "Detecting Anomalous Sensor Events in Smart Home Data for Enhancing the Living Experience," in *Artificial Intelligence and Smarter Living*, 2011, pp. 33–37, doi: https://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/viewFile/3889/4212.
[122] A. Zheng and A. Casari, *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*, First Edit. O'Reilly Media, Inc., 2018.
[123] "Wireshark · Go Deep." https://www.wireshark.org/ (accessed May 21, 2022).
[124] K. Liu, Z. Fan, M. Liu, and S. Zhang, "Hybrid Intrusion Detection Method Based on K-means and CNN for Smart Home," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2018, pp. 312–317, doi: 10.1109/CYBER.2018.8688271.
[125] M. S. Reza and J. Ma, "ICA and PCA integrated feature extraction for classification," in *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, 2016, pp. 1083–1088, doi: 10.1109/ICSP.2016.7877996.
[126] M. Nobakht, V. Sivaraman, and R. Boreli, "A Host-based Intrusion Detection and Mitigation Framework for Smart Home IoT using OpenFlow," in *2016 11th International Conference on Availability, Reliability and Security (ARES)*, 2016, pp. 147–156, doi: 10.1109/ARES.2016.64.
[127] A. Dasgupta, Y.-X. Yan, C. Ong, J.-Y. Teo, and C.-W. Lim, "Exploring Unsupervised Learning Methods for Automated Protocol Analysis," *arXiv*, 2021, doi: https://doi.org/10.48550/arXiv.2111.09061 Focus to learn more.
[128] K. Yang, S. Kpotufe, and N. Feamster, "Feature Extraction for Novelty Detection in Network Traffic," *arXiv*, 2020, doi: https://doi.org/10.48550/arXiv.2006.16993.

[129] S. Choi and J. Cho, “Novel Feature Extraction Method for Detecting Malicious MQTT Traffic Using Seq2Seq,” *Appl. Sci.*, vol. 12, no. 23, p. 12306, 2022, doi: https://doi.org/10.3390/app122312306.

[130] H. Hindy, E. Bayne, M. Bures, R. Atkinson, C. Tachtatzis, and X. Bellekens, “Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset),” in *Selected Papers from the 12th International Networking Conference*, B. Ghita and S. Shiaeles, Eds. Cham: Springer International Publishing, 2021, pp. 73–84.

[131] R. Doriguzzi-Corin, L. A. D. Knob, L. Mendozzi, D. Siracusa, and M. Savi, “Introducing Packet-Level Analysis in Programmable Data Planes to Advance Network Intrusion Detection,” *arXiv*, 2024, doi: https://doi.org/10.48550/arXiv.2307.05936.

[132] X. Li, M. Xu, P. Vijayakumar, N. Kumar, and X. Liu, “Detection of Low-Frequency and Multi-Stage Attacks in Industrial Internet of Things,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8820–8831, 2020, doi: 10.1109/TVT.2020.2995133.

[133] M. Bykova, S. Ostermann, and B. Tjaden, “Detecting network intrusions via a statistical analysis of network packet characteristics,” in *Proceedings of the Annual Southeastern Symposium on System Theory (2001)*, 2001, pp. 309–314, doi: 10.1109/SSST.2001.918537.

[134] A. Bhandari, S. Gautam, T. K. Koirala, and M. R. Islam, “Packet Sniffing and Network Traffic Analysis Using TCP—A New Approach,” in *Advances in Electronics, Communication and Computing. Lecture Notes in Electrical Engineering*, 2018, pp. 273–280, doi: https://doi.org/10.1007/978-981-10-4765-7_28.

[135] A. Ratner, S. H. Bach, H. Ehrenberg, J. Fries, S. Wu, and C. R. e, “Snorkel: Rapid training data creation with weak supervision,” *Proc. VLDB Endow.*, vol. 11, no. 3, pp. 269–282, 2017, doi: https://doi.org/10.14778/3157794.3157797.

[136] A. H. Sodhro, S. Pirbhulal, and Z. Luo, “Towards an optimal resource management for IoT based Green and sustainable smart cities,” *J. Clean. Prod.*, vol. 220, pp. 1167–1179, 2019, doi: https://doi.org/10.1016/j.jclepro.2019.01.188.

[137] Elhassan At, Aljourf M, Al-Mohanna F, and Shoukri M, “Classification of Imbalance Data using Tomek Link (T-Link) Combined with Random Under-sampling (RUS) as a Data Reduction Method,” *Glob. J. Technol. Optim.*, pp. 2–11, 2017, doi: http://dx.doi.org/10.21767/2472-1956.100011.

[138] A. H. Butt, Z. Khan, A. Khan, H. Ghazanfar, R. Zgheib, and F. Kamalov, “Performance of Sampling Methods on Imbalanced Data: Comparative Analysis,” in *2024 Advances in Science and Engineering Technology International Conferences (ASET)*, 2024, pp. 1–6, doi: 10.1109/ASET60340.2024.10708760.

[139] A. Telikani, J. Shen, J. Yang, and P. Wang, “Industrial IoT Intrusion Detection via Evolutionary Cost-Sensitive Learning and Fog Computing,” *IEEE Internet Things J.*, vol. 9, no. 22, pp. 23260–23271, 2022, doi: 10.1109/JIOT.2022.3188224.

[140] N. Sarwar, I. S. Bajwa, M. Z. Hussain, M. Ibrahim, and K. Saleem, “IoT Network Anomaly Detection in Smart Homes Using Machine Learning,” *IEEE Access*, vol. 11, pp. 119462–119480, 2023, doi: 0.1109/ACCESS.2023.3325929.
[141] A. Derhab, A. Aldweesh, A. Z. Emam, and F. A. Khan, “Intrusion Detection System for Internet of Things Based on Temporal Convolution Neural Network and Efficient Feature Engineering,” *Wirel. Commun. Mob. Comput.*, vol. 2020, no. April, pp. 1–16, 2020, doi: 10.1155/2020/6689134.
[142] Y. Sun *et al.*, “Borderline SMOTE Algorithm and Feature Selection-Based Network Anomalies Detection Strategy,” *Energies*, vol. 15, no. 13, p. 4751, 2022, doi: https://doi.org/10.3390/en15134751.
[143] A. Kumar, N. Saxena, S. Jung, and B. J. Choi, “Improving Detection of False Data Injection Attacks Using Machine Learning with Feature Selection and Oversampling,” *Energies*, vol. 15, no. 1, p. 212, 2022, doi: https://doi.org/10.3390/en15010212.
[144] A. Salehpour, M. Norouzi, M. A. Balafar, and K. SamadZamini, “A cloud-based hybrid intrusion detection framework using XGBoost and ADASYN-Augmented random forest for IoMT,” *IET Commun.*, vol. 18, pp. 1371–1390, 2024, doi: https://doi.org/10.1049/cmu2.12833.
[145] M. A. Talukder *et al.*, “Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction,” *J. Big Data*, vol. 11, no. 33, 2024, doi: https://doi.org/10.1186/s40537-024-00886-w.
[146] J. L. Leevy, T. M. Khoshgoftaar, and J. Hancock, “Using Random Undersampling and Ensemble Feature Selection for IoT Attack Prediction,” *Int. J. Reliab. Qual. Saf. Eng.*, vol. 31, no. 1, p. 2350012, 2024, doi: https://doi.org/10.1142/S0218539323500122.
[147] N. El Kamel, M. Eddabbah, Y. Lmoumen, and R. Touahni, “A Smart Agent Design for Cyber Security Based on Honeypot and Machine Learning,” *Secur. Commun. Networks*, vol. 2020, p. 9, 2020, doi: https://doi.org/10.1155/2020/8865474.
[148] R. Alasmari and A. A. Alhogail, “Protecting Smart-Home IoT Devices From MQTT Attacks: An Empirical Study of ML-Based IDS,” *IEEE Access*, vol. 12, pp. 25993–26004, 2024, doi: 10.1109/ACCESS.2024.3367113.
[149] I. S. Thaseen, V. Mohanraj, S. Ramachandran, K. Sanapala, and S.-S. Yeo, “A Hadoop Based Framework Integrating Machine Learning Classifiers for Anomaly Detection in the Internet of Things,” *Electronics*, vol. 10, no. 16, p. 1955, 2021, doi: https://doi.org/10.3390/electronics10161955.
[150] L. Liu, S. Peng, and Z. Wu, “Detection of CIFA using SMOTEBoost and LSTM in NDN,” *Comput. Secur.*, vol. 150, p. 104251, 2025, doi: https://doi.org/10.1016/j.cose.2024.104251.
[151] R. Malhotra and J. Jain, “Handling Imbalanced Data using Ensemble Learning in Software Defect Predictio,” in *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2020, pp. 300–304, doi: 10.1109/Confluence47617.2020.9058124.
[152] A. R. Javed *et al.*, “Automated cognitive health assessment in smart homes

using machine learning," *Sustain. Cities Soc.*, vol. 65, p. 102572, 2021, doi: 10.1016/j.scs.2020.102572.
[153] Ö. Sen, C. Eze, A. Ulbig, and A. Monti, "On Holistic Multi-Step Cyberattack Detection via a Graph-based Correlation Approach," 2022, doi: 10.1109/SmartGridComm52983.2022.9961016.
[154] S. Thaseen and C. A. Kumar, "Intrusion Detection Model Using fusion of Chi-square feature selection and multi class SVM," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 4, pp. 462–472, 2017, doi: 10.1016/j.jksuci.2015.12.004.
[155] M. Ozkan-Okay, R. Samet, Ö. Aslan, S. Kosunalp, T. Iliev, and I. Stoyanov, "A Novel Feature Selection Approach to Classify Intrusion Attacks in Network Communications," *Appl. Sci.*, vol. 13, no. 19, p. The fast development of communication technologies, 2023, doi: https://doi.org/10.3390/app131911067.
[156] K. Albulayhi, Q. A. Al-Haija, S. A. Alsuhibany, A. A. Jillepalli, M. Ashrafuzzaman, and F. T. Sheldon, "IoT Intrusion Detection Using Machine Learning with a Novel High Performing Feature Selection Method," *Appl. Sci.*, vol. 12, no. 5015, 2022, doi: https://doi.org/10.3390/app12105015.
[157] Z. R. S. Elsi *et al.*, "Feature Selection using Chi Square to Improve Attack Detection Classification in IoT Network: Work in Progress," in *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI2022)*, 2022, pp. 226–232, doi: 10.23919/EECSI56542.2022.9946621.
[158] 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," in *International Conference on Advanced Information Networking and Applications*, 2020, pp. 458–469, doi: 10.1007/978-3-030-15032-7_39.
[159] M. Al-Sarem, F. Saeed, E. H. Alkhammash, and N. S. Alghamdi, "An Aggregated Mutual Information Based Feature Selection with Machine Learning Methods for Enhancing IoT Botnet Attack Detection," *Sensors*, vol. 22, p. 185, 2021, doi: https://doi.org/10.3390/s22010185.
[160] M. A. F. A. Fida, T. Ahmad, and M. Ntahobari, "Variance Threshold as Early Screening to Boruta Feature Selection for Intrusion Detection System," in *13th International Conference on Information & Communication Technology and System (ICTS)*, 2021, pp. 46–50, doi: 10.1109/ICTS52701.2021.9608852.
[161] A. A. Qasem, M. H. Qutqut, F. Alhaj, and A. Kitana, "SRFE: A stepwise recursive feature elimination approach for network intrusion detection systems," *Peer-to-Peer Netw. Appl.*, vol. 12, 2024, [Online]. Available: https://link.springer.com/article/10.1007/s12083-024-01763-2.
[162] A. S. Kumar, T. J. Nagalakshmi, and M. N, "Improving the Accuracy of Intrusion Detection System in the Detection of DoS using Naive Bayes with Lasso Feature Elimination and Comparing with Naive Bayes without Feature Elimination in Wireless Adhoc Network," in *International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*,

2023, pp. 1–7, doi: 10.1109/ACCAI58221.2023.10199248.
[163] G. Lazrek, K. Chetioui, Y. Balboul, S. Mazer, and M. El Bekkali, “An RFE/Ridge-ML/DL based anomaly intrusion detection approach for securing IoMT system,” *Results Eng.*, vol. 23, 2024, doi: https://doi.org/10.1016/j.rineng.2024.102659.
[164] A. Srinivasan and P. Deepalakshmi, “ENetRM: ElasticNet Regression Model based malicious cyber-attacks prediction in real-time server,” *Meas. Sensors*, vol. 25, 2023, doi: https://doi.org/10.1016/j.measen.2022.100654.
[165] J. Ashraf, G. M. Raza, B.-S. Kim, A. W. Kim, and Hye-Young, “Making a Real-Time IoT Network Intrusion-Detection System (INIDS) Using a Realistic BoT–IoT Dataset with Multiple Machine-Learning Classifiers,” *Appl. Sci.*, vol. 15, no. 2043, pp. 1–19, 2025, doi: https://doi.org/10.3390/app15042043.
[166] H. Hindy, C. Tachtatzis, R. Atkinson, E. Bayne, and X. Bellekens, “MQTT-IoT-IDS2020: MQTT Internet of Things Intrusion Detection Dataset,” *IEEE Dataport*, 2020, doi: http://10.0.82.235/bhxy-ep04.
[167] Z. R. S. Elsi, D. Stiawan3, B. Y. Suprapto2, M. A. S. Arifin4, M. Y. Idris5, and R. Budiarto, “Enhanced Intrusion Detection in IoT Smart Homes: Leveraging Binary and Multi-Class Classification Models,” *Int. J. Online Biomed. Eng.*, vol. 21, no. 5, pp. 63–86, 2025, doi: https://doi.org/10.3991/ijoe.v21i05.53485.
[168] A. J. Hintaw, S. Manickam, M. F. Aboalmaaly, and S. Karuppayah, “MQTT Vulnerabilities, Attack Vectors and Solutions in the Internet of Things (IoT),” *IETE J. Res.*, pp. 3368–3397, 2021, doi: https://doi.org/10.1080/03772063.2021.1912651.
[169] C. Patel and N. Doshi, “A Novel MQTT Security framework In Generic IoT Model,” *Procedia Comput. Sci.*, vol. 171, pp. 1399–1408, 2020, doi: 10.1016/j.procs.2020.04.150.
[170] I. Makhdoom, M. Abolhasan, J. Lipman, R. P. Liu, and W. Ni, “Anatomy of Threats to the Internet of Things,” *IEEE Commun. Surv. Tutorials*, vol. 21, no. 2, pp. 1636–1675, 2019, doi: 10.1109/COMST.2018.2874978.
[171] F. Chen, Y. Huo, J. Zhu, and D. Fan, “A Review on the Study on MQTT Security Challenge,” in *2020 IEEE International Conference on Smart Cloud, SmartCloud 2020 (2020)*, 2020, pp. 128–133, doi: 10.1109/SmartCloud49737.2020.00032.
[172] H. A. P and K. K., “Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things,” *J. Wirel. Commun. Netw.*, vol. 90, 2019, doi: https://doi.org/10.1186/s13638-019-1402-8.
[173] J. Arshad, M. A. Azad, R. Amad, K. Salah, M. Alazab, and R. Iqbal, “A Review of Performance, Energy and Privacy of Intrusion Detection Systems for IoT,” *Electronics*, vol. 9, no. 4, p. 629, 2020, doi: https://doi.org/10.3390/electronics9040629.
[174] S. Mishra and A. Paul, “A Critical Analysis of Attack Detection Schemes in IoT and Open Challenges,” in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 2020, pp.

57–62, doi: 10.1109/GUCON48875.2020.9231077.
[175] S. J. Saidi *et al.*, “A Haystack Full of Needles: Scalable Detection of IoT Devices in the Wild,” in *IMC ’20: Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 87–100, doi: https://doi.org/10.1145/3419394.3423650.
[176] S. Bhanja and A. Das, “Impact of Data Normalization on Deep Neural Network for Time Series Forecasting,” *arXiv*, 2018, doi: https://doi.org/10.48550/arXiv.1812.05519.
[177] E. Blessing and H. Klaus, “Normalization and Standardization: Methods to preprocess data to have consistent scales and distributions,” vol. 2237. p. 10, 2023, [Online]. Available: https://www.researchgate.net/publication/377123133_Normalization_and_Standardization_Methods_to_preprocess_data_to_have_consistent_scales_and_distributions.
[178] H. Alfares and O. Banimelhem, “Comparative Analysis of Machine Learning Techniques for Handling Imbalance in IoT-23 Dataset for Intrusion Detection Systems,” in *2024 11th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, 2024, pp. 112–119, doi: 10.1109/IOTSMS62296.2024.10710296.
[179] K.-A. Tait, J. S. Khan, F. Alqahtani, A. A. Shah, F. A. Khan, and M. U. Rehman, “Intrusion Detection using Machine Learning Techniques: An Experimental Comparison,” in *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1–10, doi: 10.1109/ICOTEN52080.2021.9493543.
[180] B. Yan and G. Han, “LA-GRU: Building Combined Intrusion Detection Model Based on Imbalanced Learning and Gated Recurrent Unit Neural Network,” *Secur. Commun. Networks*, pp. 1–13, 2018, doi: https://doi.org/10.1155/2018/6026878.
[181] A. EL HARIRI, M. MOUITI, O. HABIBI, and M. LAZAAR, “Improving Deep Learning Performance Using Sampling Techniques for IoT Imbalanced Data,” *Procedia Comput. Sci.*, vol. 224, pp. 180–187, 2023, doi: https://doi.org/10.1016/j.procs.2023.09.026.
[182] J. Jiang *et al.*, “A dynamic ensemble algorithm for anomaly detection in IoT imbalanced data streams,” *Comput. Commun.*, vol. 194, pp. 250–257, 2022, doi: https://doi.org/10.1016/j.comcom.2022.07.034.
[183] T. Riston *et al.*, “Oversampling Methods for Handling Imbalance Data in Binary Classification,” *Comput. Sci. Its Appl. – ICCSA 2023 Work. ICCSA 2023*, pp. 3–23, 2023, doi: https://doi.org/10.1007/978-3-031-37108-0_1.
[184] L. Xue and T. Zhu, “Hybrid resampling and weighted majority voting for multi-class anomaly detection on imbalanced malware and network traffic data,” *Eng. Appl. Artif. Intell.*, vol. 128, p. 107568, 2024, doi: https://doi.org/10.1016/j.engappai.2023.107568.
[185] B. S. Sharmila and R. Nagapadma, “RT-IoT2022 [Dataset].” UCI Machine Learning Repository, 2023, doi: 10.24432/C5P338.
[186] S. Bagui and K. L, “Resampling imbalanced data for network intrusion detection datasets,” *J. Big Data*, vol. 8, no. 6, 2021, doi:

https://doi.org/10.1186/s40537-020-00390-x.
[187] A. Vaishnavi, B. R. Ganesh, A. D. Reddy, and K. L. Kumar, "Ensemble-Learning-Based Deep Neural Network Attack Classification of Imbalanced IoT Intrusion Data," *Int. J. Inf. Technol. Comput. Eng.*, vol. 12, no. 3, pp. 634–646, 2024, [Online]. Available: https://ijitce.org/index.php/ijitce/article/view/715.
[188] J. Song, X. Wang, M. He, and L. Jin, "CSK-CNN: Network Intrusion Detection Model Based on Two-Layer CNN for Handling Imbalanced Dataset," *Information*, vol. 14, no. 2, p. 130, 2023, doi: https://doi.org/10.3390/info14020130.
[189] C. Zhang, K. C. Tan, H. Li, and G. S. Hong, "A Cost-Sensitive Deep Belief Network for Imbalanced Classification," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 1, pp. 109–122, 2019, doi: 10.1109/TNNLS.2018.2832648.
[190] H. Zhang, L. Huang, C. Q. Wu, and Z. Li, "An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset," *Comput. networks*, vol. 177, p. 107315, 2020, doi: https://doi.org/10.1016/j.comnet.2020.107315.
[191] Z. Fan, S. Sohail, F. Sabrina, and X. Gu, "Sampling-Based Machine Learning Models for Intrusion Detection in Imbalanced Dataset," *Electronics*, vol. 13, no. 10, p. 1878, 2024, doi: https://doi.org/10.3390/electronics13101878.
[192] M. Y. Arafat, S. Hoque, S. Xu, and D. M. Farid, "Machine learning for mining imbalanced data," *IAENG Int. J. Comput. Sci.*, vol. 46, no. 2, pp. 332–348, 2019, [Online]. Available: https://www.iaeng.org/IJCS/issues_v46/issue_2/IJCS_46_2_21.pdf.
[193] B. Weinger, J. Kim, A. Sim, M. Nakashima, N. Moustafa, and K. J. Wu, "Enhancing IoT anomaly detection performance for federated learning," *Digit. Commun. Networks*, vol. 8, no. 3, pp. 314–323, 2022, doi: https://doi.org/10.1016/j.dcan.2022.02.007.
[194] O. D. Okey *et al.*, "BoostedEnML: Efficient Technique for Detecting Cyberattacks in IoT Systems Using Boosted Ensemble Machine Learning," *Sensors*, vol. 22, no. 19, p. 7409, 2022, doi: https://doi.org/10.3390/s22197409.
[195] M. Sarhan, S. Layeghy, and M. Portmann, "Feature Analysis for Machine Learning-based IoT Intrusion Detection," *ArXiv*, 2022, doi: https://doi.org/10.48550/arXiv.2108.12732 Focus to learn more.
[196] R. Samdekar, D. S. M. . Ghosh, and K. Srinivas, "Efficiency Enhancement of Intrusion Detection in Iot Based on Machine Learning Through Bioinspire," in *Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2021, pp. 383–387, doi: 10.1109/ICICV50876.2021.9388392.
[197] A. Alabrah, "A Novel Study: GAN-Based Minority Class Balancing and Machine-Learning-Based Network Intruder Detection Using Chi-Square Feature Selection," *Appl. Sci.*, vol. 12, no. 22, p. 11662, 2022, doi: https://doi.org/10.3390/app122211662.

[198] S. Dwivedi, M. Vardhan, and S. Tripathi, “Distributed Denial-of-Service Prediction on IoT Framework by Learning Techniques,” *Open Comput. Sci.*, vol. 10, pp. 220–230, 2020, doi: https://doi.org/10.1515/comp-2020-0009.

[199] S. Wijethilaka and M. Liyanage, “Realizing Internet of Things with Network Slicing: Opportunities and Challenges,” no. January, pp. 1–6, 2021, doi: 10.1109/ccnc49032.2021.9369637.

[200] S. Kaushik, A. Bhardwaj, A. Alomari, S. Bharany, A. Alsirhani, and M. M. Alsha, “Efficient, Lightweight Cyber Intrusion Detection System for IoT Ecosystems Using MI2G Algorithm,” *computers*, vol. 11, no. 142, 2022, doi: https://doi.org/10.3390/computers11100142.

[201] S. S. Kareem, R. R. Mostafa, F. A. Hashim, and H. M. El-Bakry, “An Effective Feature Selection Model Using Hybrid Metaheuristic Algorithms for IoT Intrusion Detection,” *Sensors*, vol. 22, no. 1396, 2022, doi: https://doi.org/10.3390/s22041396.