PENERAPAN *MULTI-CLASSIFICATION* SERANGAN SIBER DENGAN METODE *NAÏVE BAYES* DAN *CHI-SQUARE FEATURE SELECTION*

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

OLEH:

NURUL FITRIA

09011282025042

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2024

HALAMAN PENGESAHAN

PENERAPAN *MULTI-CLASSIFICATION* SERANGAN SIBER DENGAN METODE *NAÏVE BAYES* DAN *CHI-SQUARE FEATURE SELECTION*

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh
Nurul Fitria
09011282025042

Indralaya, 02 Mei 2024

Mengetahui,

Ketua Jurusan Sistem Komputer	Pembimbing,
Dr. Ir. Sukemi, M.T.	**Ahmad Heryanto, S. Kom., M.T.**
NIP. 196612032006041001	**NIP. 198701222015041002**

VALIDITY SHEET

IMPLEMENTATION OF MULTI-CLASSIFICATION CYBER ATTACKS WITH NAÏVE BAYES METHOD AND CHI-SQUARE FEATURE SELECTION

THESIS

Submitted to Complete One of the Requirements Earned a Bachelor's Degree in Computer Science

By

Nurul Fitria

09011282025042

Indralaya, 24/5 2024

Acknowledged,

Head of Computer System Department	Supervisor
Dr. Ir. Sukemi, M.T.	**Ahmad Heryanto, S. Kom., M.T.**
NIP. 196612032006041001	**NIP. 198701222015041002**

Penerapan *Multi-Classification* Serangan Siber dengan Metode *Naïve Bayes* dan *Chi-Square Feature Selection*

NURUL FITRIA (09011282025042)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
Email : fitriahnurul834@gmail.com

Abstrak

Metode *Naïve Bayes* telah terbukti efektif dalam mendeteksi, mengevaluasi, dan melakukan *multi-classification* serangan siber. Dengan memanfaatkan algoritma *Chi-Square Feature Selection*, akurasi dapat ditingkatkan dengan pemilihan fitur-fitur yang paling relevan untuk mendeteksi serangan, memungkinkan klasifikasi serangan berdasarkan jenisnya. Dataset yang digunakan dalam pengujian dan *multi-classification* meliputi CIC-IDS2018, CIC-IDS2017, ISCX2012, dan KDD-CUP 1999 dalam format CSV. Penerapan Chi-Square Feature Selection berhasil mendapatkan nilai F1-Score yang tinggi di setiap dataset, pemilihan F1-Score ini berdasarkan dataset yang digunakan tidak seimbang. Dimana diperoleh hasil yang optimal pada rasio 80:20 untuk dataset CIC-IDS2017 dan 2018, rasio 20:80 untuk dataset ISCX 2012, dan rasio 70:30 untuk dataset KDD-CUP 1999. Selain pemilihan rasio yang tepat, pemilihan model juga memiliki peran penting, dimana model *Gaussian Naïve Bayes* efektif untuk dataset CIC-IDS2018 dan ISCX 2012, sementara model *Multinomial Naïve Bayes* lebih unggul untuk dataset CIC-IDS2017 dan KDD CUP 1999. Validasi menegaskan pentingnya memilih model *Naïve Bayes* yang sesuai dengan karakteristik dataset, menjadi kunci dalam mencapai performa deteksi serangan siber yang optimal. Dengan pendekatan yang cermat terhadap pemilihan fitur, rasio data, dan model, sistem dapat menghasilkan hasil yang akurat dan efisien dalam mendeteksi serangan siber.

Kata Kunci: *Naïve Bayes*, *Chi-Square feature selection*, *multi-classification*, deteksi serangan siber.

Mengetahui,

Ketua Jurusan Sistem Komputer	**Pembimbing Tugas Akhir,**
Dr. Ir. Sukemi, M.T.	**Ahmad Heryanto, S.Kom., M.T.**
NIP. 196612032006041001	**NIP. 198701222015041002**

Implementation Of Multi-Classification Cyber Attacks with Naïve Bayes Method and Chi-Square Feature Selection

NURUL FITRIA (09011282025042)
Department of Computer Systems, Faculty of Computer Science, Sriwijaya University
Email : fitriahnurul834@gmail.com

Abstract

The Naïve Bayes method has proven effective in detecting, evaluating, and performing multi-classification of cyberattacks. By leveraging the Chi-Square Feature Selection algorithm, accuracy can be improved by selecting the most relevant features for detecting attacks, enabling classification based on their types. The datasets used for testing and multi-classification include CIC-IDS2018, CIC-IDS2017, ISCX2012, and KDD-CUP 1999 in CSV format. The implementation of Chi-Square Feature Selection successfully achieved high F1-Score values on each dataset, considering the datasets are imbalanced. Optimal results were obtained with an 80:20 ratio for CIC-IDS2017 and 2018 datasets, a 20:80 ratio for ISCX 2012 dataset, and a 70:30 ratio for KDD-CUP 1999 dataset. Besides the appropriate ratio selection, the choice of model also plays a crucial role, where the Gaussian Naïve Bayes model is effective for CIC-IDS2018 and ISCX 2012 datasets, while the Multinomial Naïve Bayes model performs better for CIC-IDS2017 and KDD-CUP 1999 datasets. Validation confirms the importance of selecting a Naïve Bayes model that matches the dataset characteristics, key to achieving optimal performance in cyberattack detection. With a careful approach to feature selection, data ratio, and model choice, the system can produce accurate and efficient results in detecting cyberattacks.

Keywords: Naïve Bayes, Chi-Square feature selection, multi-classification, cyber-attack detection.

Acknowledged,

Head of Computer System Department	**Supervisor**
Dr. Ir. Sukemi, M.T.	**Ahmad Heryanto, S. Kom., M.T.**
NIP. 196612032006041001	**NIP. 198701222015041002**

HALAMAN PERSETUJUAN

Telah Diuji dan lulus pada

Hari : Kamis

Tanggal : 04 April 2024

Tim Penguji :

1. Ketua : Huda Ubaya, M. T.

2. Sekretaris : Abdurahman, S.Kom., M.Han

3. Penguji : Dr. Rossi Passarella, M.Eng.

4. Pembimbing : Ahmad Heryanto, S. Kom., M. T

Digitally signed by Rossi Passarella
DN: cn=Rossi Passarella, c=ID, o=Universitas Sriwijaya, ou=Jurusan Sistem Komputer, email=passarella.rossi@unsri.ac.id
Reason: I am approving this document
Location: Shah Alam, Malaysia
Date: 2024.04.27 07:43:55 +07'00'

Mengetahui,
Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama	:	Nurul Fitria
NIM	:	09011282025042
Judul	:	Penerapan *Multi-Classification* Serangan Siber dengan Metode *Naïve Bayes* dan *Chi Square Feature Selection*

Hasil Pengecekan Software Ithenticate/ Turnitin: 2%

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

Indralaya, 02 Mei 2024

Nurul Fitria
09011282025042

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji dan syukur atas kehadirat Allah SWT, karena berkat Rahmat dan Karunia-Nya lah sehingga penulis dapat menyelesaikan penyusunan proposal tugas akhir ini dengan judul **"Penerapan *Multi-Classification* Serangan Siber Dengan Metode *Naïve Bayes* Dan *Chi-Square Feature Selection*"**.

Pada kesempatan proposal tugas akhir ini tidak terlepas dari bimbingan dan bantuan dari berbagai pihak, sehingga proposal tugas akhir ini dapat diselesaikan dengan baik. Oleh karena itu, dalam kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah Subhanahu Wata'ala, yang telah memberikan berkah serta nikmat kesehatan dan kemudahan kepada penulis.
2. Kedua Orang Tua, Keluarga dan Teman-teman yang selalu mendoakan serta memberikan motivasi dan semangat.
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 Prof. Dr. Ir. Bambang Tutuko, M.T. selaku Dosen Pembimbing Akademik penulis.
6. Bapak Ahmad Heryanto, S. Kom., M.T. Selaku Dosen Pembimbing Skripsi yang telah meluangkan waktunya untuk membimbing serta memberikan saran dan motivasi terbaik kepada penulis dalam menyelesaikan Tugas Akhir.
7. Bapak Yopi selaku Admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal-hal administrasi.
8. Teman-teman seperjuangan Jurusan Sistem Komputer Angkatan 2020, terutama Rizki Elinda Sari, Zuli Yanti, Septiani Kusuma Ningrum, Vijiantika Fajaria Sastri dan Arinda Intan Safitri yang telah membantu dan menjadi *support system* penulis.
9. Dan seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan bantuan, semangat serta doa.

Dalam penulisan skripsi ini, penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan, maka dari itu penulis sangat mengharapkan kritik dan saran yang bersifat membangun agar lebih baik lagi di kemudian hari dari semua pihak yang berkenan.

Akhir kata dengan segala keterbatasan, penulis berharap semoga laporan skripsi ini menghasilkan sesuatu yang bermanfaat bagi kita semua khususnya bagi mahasiswa/i Jurusan Sistem Komputer Universitas Sriwijaya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, 02 Mei 2024

Penulis,

Nurul Fitria

NIM. 09011282025042

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR S*YNTAX*

BAB I
PENDAHULUAN

1.1 Latar Belakang

Dengan semakin banyaknya perangkat yang terhubung ke internet maka akan semakin meningkat pula serangan siber yang akan dihadapi [1]. Dimana, serangan siber telah menjadi ancaman yang signifikan bagi perusahaan besar maupun kecil serta seluruh masyarakat. Dalam beberapa tahun terakhir, serangan siber semakin merajalela dan terus meningkat baik dalam jumlah maupun jenis serangan yang dilakukan [2]. Selain itu, serangan siber ini bukan hanya mengancam keamanan sistem informasi, tetapi juga dapat merusak reputasi perusahaan dan menimbulkan kerugian secara finansial [1]. Ancaman ini datang dari berbagai pihak, baik dari peretas yang ingin mengambil keuntungan dengan mencuri informasi penting [2]. Oleh karena itu, deteksi serangan siber menjadi sangat penting dalam mengamankan sistem informasi.

Dalam konteks deteksi serangan siber, klasifikasi merupakan proses mengelompokkan suatu aktivitas pada sistem komputer dalam kategori yang berbeda berdasarkan karakteristik nya. Adapun dalam deteksi serangan siber memiliki 2 jenis klasifikasi yaitu klasifikasi biner dan klasifikasi multiclass. Klasifikasi biner yaitu membagi data menjadi dua kelas saja yang saling eksklusif, yang biasa lebih dikenal dengan serangan (positif) dan benign (negatif). Sedangkan klasifikasi Multiclass ialah membagi data menjadi lebih dari dua kelas atau dua jenis serangan yang berbeda-beda, seperti serangan DoS (*Denial of Service*), serangan SQL Injection, serangan Malware, dan serangan lainnya. Adapun kelebihan dari klasifikasi multiclass dibandingkan klasifikasi binary yaitu klasifikasi multiclass lebih informatif karena dapat mengidentifikasi lebih banyak jenis serangan secara spesifik[3].

Dalam sistem deteksi multi-klasifikasi serangan siber, metode Naïve Bayes (NB) dapat digunakan untuk mengklasifikasikan data log aktivitas jaringan ke dalam beberapa kelas serangan siber yang berbeda bahkan mampu mengklasifikasikan nya menjadi kategori serangan atau bukan serangan [4]. Metode ini dapat menghasilkan model klasifikasi yang efektif dan efisien karena

mampu mengatasi masalah dimensi data yang tinggi dan pengaruh interaksi antara fitur yang tidak signifikan. Sehingga dapat membantu mempercepat proses identifikasi serangan siber yang terjadi dan mencegah kerugian yang lebih besar. Metode NB merupakan salah satu teknik klasifikasi yang paling sederhana, efektif, efisien dan populer dalam mengidentifikasi dan mencegah serangan siber. Selain itu, metode ini juga relatif mudah untuk digunakan dan hanya memerlukan waktu pelatihan yang singkat [5][6]. Hal inilah yang membuat metode NB menjadi pilihan yang baik dan cocok untuk digunakan dalam sistem deteksi serangan siber [6].

Metode NB didasarkan pada teorem bayes, yang menyatakan bahwa untuk menghitung probabilitas kelas pada fitur atau atribut yang digunakan, berdasarkan probabilitas kondisional dari hipotesis tersebut dan data yang diamati [5][7]. Dalam mendeteksi serangan siber metode NB akan digunakan untuk mengklasifikasi aktivitas jaringan menjadi beberapa kelas, seperti normal, serangan DoS, Serangan probing, dan serangan lainnya [5]. Salah satu algoritma yang bisa digunakan dalam sistem deteksi serangan siber adalah algoritma *feature selection*. Algoritma ini berguna untuk memilih fitur yang paling relevan untuk tujuan klasifikasi, sehingga mengurangi dimensi data serta meningkatkan kinerja dari metode NB [6].

Adapun dataset yang akan diterapkan pada multi-klasifikasi serangan siber ini yaitu dataset CIC-IDS 2018, dataset CIC-IDS 2017, dataset ISCX 2012 dan dataset KDD-CUP 1999. Tujuan lain dari penelitian ini yaitu dapat membandingkan dataset mana yang paling cocok dengan metode NB ini, dikarenakan metode NB ini kurang cocok untuk dataset yang memiliki dengan nilai kontinu atau atribut yang sangat berkorelasi satu sama lain. Oleh karena itu sebelum menerapkan metode NB pada dataset, akan dilakukan analisis terlebih dahulu untuk mengetahui apakah metode ini cocok atau tidak untuk masalah klasifikasi.

Pada penelitian [8] membahas mengenai penerapan metode NB untuk instruction detection system (IDS) pada dataset NSL-KDD dengan tujuan serangan-serangan baru dapat ter klasifikasi dan mendapatkan nilai akurasi kebenaran sebesar 81-84,67 %. Pada penelitian [9], sudah dilakukan klasifikasi menggunakan dataset standar NSL-KDD, UNSW-NB15, dan CIC-IDS2017 dan nilai akurasi yang didapatkan mencapai 97% pada dataset NSL-KDD, 86,9% pada dataset UNSWNB15, dan 98,59% pada dataset CIC-IDS2017. Namun model yang diusulkan tidak berfungsi dengan baik dalam klasifikasi multikelas, meskipun

mencapai akurasi keseluruhan sebesar 97% dan memiliki tingkat kesalahan positif palsu yang rendah. Sedangkan pada penelitian[10], menggunakan metode *Chi-Square feature selection* dan mendapatkan nilai akurasi yang tinggi, hal ini menunjukkan bahwa fitur yang diusulkan cukup diskriminatif untuk mencapai tujuannya. Pada penelitian tersebut mendapatkan nilai akurasi tertinggi sebesar 98.1% dengan *Support Vector Machine*. Oleh karena itu, mengacu dari penelitian sebelumnya, maka pada penelitian ini akan menggunakan metode *Chi-Square feature selectio*n yang berfungsi untuk memilih fitur-fitur yang paling relevan dalam mendeteksi serangan siber dengan ambang batas nilai p (p-value) sebesar 0,05 untuk mencapai tujuan. Hal ini digunakan untuk memastikan kemandirian data, dimana nantinya metode *Chi-Square* akan menghitung perbedaan antara jumlah pengamatan (O) dan jumlah yang diharapkan (E) dalam dataset [11].

Dengan uraian di atas, penulis memilih judul **"Penerapan *Multi-Classification* Serangan Siber Dengan Metode Naïve Bayes Dan *Chi-Square Feature Selection*".** Hasil dari penelitian tugas akhir ini nantinya, diharapkan dapat membantu serta memberi wawasan bagi yang membutuhkan dan juga dapat menjadi salah satu cara untuk mengatasi serangan siber saat ini.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, peeliti merumuskan masalah yaitu sebagai berikut:

1. Bagaimana menerapkan metode NB agar dapat digunakan untuk mendeteksi dan melakukan *multi-classification* serangan siber?
2. Bagaimana metode NB dan algoritma *Chi-Square feature selection* dapat diimplementasikan pada dataset yang akan dianalisis?
3. Bagaimana performa metode NB dapat dianalisis dan dievaluasi dalam konteks deteksi dan *multi-classification* serangan siber?

1.3 Tujuan

Berdasarkan rumusan masalah di atas, penelitian ini memiliki beberapa tujuan yang harus dicapai yaitu:

1. Menerapkan metode NB pada sistem deteksi *multi-classification* serangan siber terhadap dataset yang diterapkan.
2. Menerapkan algoritma *Chi-Square feature selection* untuk memilih fitur-fitur yang paling relevan dalam mendeteksi serangan siber sehingga dapat menentukan model terbaik.
3. Mengimplementasikan metode NB untuk memultiklasifikasi beberapa serangan siber berdasarkan jenis serangannya.

1.4 Manfaat

Berdasarkan tujuan dari penelitian ini, memiliki beberapa manfaat antara lain sebagai berikut:

1. Penerapan metode NB dapat membantu mengidentifikasi jenis serangan sehingga mampu meningkatkan keakuratan dan efisiensi dalam mendeteksi serangan siber.
2. Dapat melakukan *multi-classification* pada beberapa dataset menggunakan metode NB dan *Chi-Square feature selection*.
3. Dapat memperoleh tingkat akurasi, presisi, Recall dan F1-Score dari metode NB yang didukung dengan algoritma *Chi-Square feature selection*.

1.5 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah:

1. Dataset yang digunakan pada penelitian ini menggunakan lebih dari satu dataset yaitu dataset CIC-IDS2018, CIC-IDS2017, ISCX2012 dan Dataset KDD-CUP 1999.
2. Pada penelitian ini, melakukan deteksi serangan siber dan *multi-classification* menggunakan metode NB.
3. *Multi-classification* yang dilakukan akan menggunakan algoritma *Chi-Square feature selection.*

1.6 Metode Penelitian

Pada penelitian ini nantinya akan melewati beberapa tahapan-tahapan yaitu sebagai berikut:

1. Tahapan pertama (perumusan masalah)

   Tahapan pertama adalah perumusan masalah yaitu penentuan pokok permasalahan mengenai sistem deteksi dan *multi-classification* serangan siber.

2. Tahap kedua (literatur)

   Tahap kedua adalah tahapan literatur yaitu penulis mengumpulkan referensi jurnal dan atau presiding internasional bereputasi sebanyak-banyaknya atau minimal 30 referensi yang membahas tentang sistem deteksi serta *multi-classification* maupun jurnal yang berkaitan dengan metode naïve bayes atau juga bisa dengan jenis metode lain sebagai referensi.

3. Tahap ketiga (rancang sistem)

   Tahap ketiga yaitu tahapan perancangan sistem, berdasarkan tahap pertama dan tahap kedua yang digunakan.

4. Tahap keempat (persiapan data)

   Pada tahap ini yaitu mengumpulkan dataset yang akan diuji dan diklasifikasi yaitu dataset CIC-IDS2018, CIC-IDS2017, ISCX2012 dan dataset KDD-CUP 1999 yang telah diubah ke dalam format Command-Separated Values (CSV).

5. Tahap kelima (pengujian dan klasifikasi)

   Tahap kelima yaitu tahapan lanjutan dari tahapan keempat yang telah diselesaikan. Dengan melakukan sistem deteksi terhadap serangan siber kemudian melakukan *multi-classification* antara beberapa serangan siber menggunakan metode NB dan algoritma *Chi-Square feature selection*.

6. Tahap keenam (analisa)

Tahapan keenam yaitu analisa, dimana analisa data diperoleh dari proses pengujian dengan sistem deteksi serangan siber dan hasil dari *multi-classification* ke beberapa dataset..

7. Tahap ketujuh (kesimpulan dan saran)

Pada tahapan terakhir adalah membuat kesimpulan serta membuat saran yang nantinya dapat berguna untuk penulis selanjutnya yang akan dijadikan acuan.

1.7 Sistematika Penulisan

Dalam proses penyusunan laporan tugas akhir ini, penulis menerapkan sistematika penulisan agar memudahkan dalam memahami isi dari tiap-tiap bab yang disusun dalam skripsi ini.

BAB I PENDAHULUAN

Bab ini akan menjelaskan mengenai latar belakang, perumusan masalah, Tujuan, Manfaat, metode penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini mencakup penelitian terkait, landasan teori, dan tabel ringkasan studi literatur terkait dengan serangan siber.

BAB III METODOLOGI PENELITIAN

Pada bab ketiga akan membahas tentang dataset, perangkat-perangkat yang digunakan dan membuat diagram proses penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil pengujian dan menganalisis terhadap hasil penelitian yang disajikan berdasarkan metode dari setiap hasil yang diperoleh.

BAB V KESIMPULAN DAN SARAN

Bab ini akan menyajikan kesimpulan serta saran dari penelitian ini untuk pengembangan lebih lanjut kedepannya.

DAFTAR PUSTAKA

[1] M. Alawida, A. E. Omolara, O. I. Abiodun, and M. Al-Rajab, "A deeper look into cybersecurity issues in the wake of Covid-19: A survey," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8176–8206, 2022, doi: 10.1016/J.JKSUCI.2022.08.003.

[2] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments," *Energy Reports*, vol. 7, pp. 8176–8186, 2021, doi: 10.1016/J.EGYR.2021.08.126.

[3] X. Gao *et al.*, "A multiclass classification using one-versus-all approach with the differential partition sampling ensemble," *Eng. Appl. Artif. Intell.*, vol. 97, p. 104034, Jan. 2021, doi: 10.1016/j.engappai.2020.104034.

[4] R. Panigrahi *et al.*, "Intrusion detection in cyber–physical environment using hybrid Naïve Bayes—Decision table and multi-objective evolutionary feature selection," *Comput. Commun.*, vol. 188, pp. 133–144, Apr. 2022, doi: 10.1016/j.comcom.2022.03.009.

[5] S. ur Rehman *et al.*, "DIDDOS: An approach for detection and identification of Distributed Denial of Service (DDoS) cyberattacks using Gated Recurrent Units (GRU)," *Futur. Gener. Comput. Syst.*, vol. 118, pp. 453–466, 2021, doi: 10.1016/J.FUTURE.2021.01.022.

[6] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5432–5435, 2009, doi: 10.1016/J.ESWA.2008.06.054.

[7] D. M. Farid, L. Zhang, C. M. Rahman, M. A. Hossain, and R. Strachan, "Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1937–1946, 2014, doi: 10.1016/J.ESWA.2013.08.089.

[8] A. Prasetyo, L. Affandi, and D. Arpandi, "Implementasi Metode Naive Bayes Untuk Intrusion Detection System (Ids)," *J. Inform. Polinema*, vol. 4, no. 4, p. 280, 2018, doi: 10.33795/jip.v4i4.220.

[9] M. Vishwakarma and N. Kesswani, "A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic

envelop method for anomaly detection," *Decis. Anal. J.*, vol. 7, no. April, p. 100233, Jun. 2023, doi: 10.1016/j.dajour.2023.100233.

[10] I. Sumaiya Thaseen and C. Aswani 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, Oct. 2017, doi: 10.1016/j.jksuci.2015.12.004.

[11] L. A. C. Ahakonye, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "SCADA intrusion detection scheme exploiting the fusion of modified decision tree and Chi-square feature selection," *Internet of Things*, vol. 21, p. 100676, Apr. 2023, doi: 10.1016/j.iot.2022.100676.

[12] S. M. Kasongo, "A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework," *Comput. Commun.*, vol. 199, no. January 2021, pp. 113–125, 2023, doi: 10.1016/j.comcom.2022.12.010.

[13] W. Ding, M. Abdel-Basset, and R. Mohamed, "DeepAK-IoT: An effective deep learning model for cyberattack detection in IoT networks," *Inf. Sci. (Ny).*, vol. 634, no. January, pp. 157–171, Jul. 2023, doi: 10.1016/j.ins.2023.03.052.

[14] M. M. Alani, L. Mauri, and E. Damiani, "A two-stage cyber attack detection and classification system for smart grids," *Internet of Things*, vol. 24, p. 100926, Dec. 2023, doi: 10.1016/j.iot.2023.100926.

[15] P. R. Kanna and P. Santhi, "Hybrid Intrusion Detection using MapReduce based Black Widow Optimized Convolutional Long Short-Term Memory Neural Networks," *Expert Syst. Appl.*, vol. 194, p. 116545, May 2022, doi: 10.1016/j.eswa.2022.116545.

[16] Z. Wang, Y. Liu, D. He, and S. Chan, "Intrusion detection methods based on integrated deep learning model," *Comput. Secur.*, vol. 103, 2021, doi: 10.1016/j.cose.2021.102177.

[17] T. Thilagam and R. Aruna, "Intrusion detection for network based cloud computing by custom RC-NN and optimization," *ICT Express*, vol. 7, no. 4, pp. 512–520, Dec. 2021, doi: 10.1016/j.icte.2021.04.006.

[18] S. Kabir, S. Sakib, M. A. Hossain, S. Islam, and M. I. Hossain, "A

Convolutional Neural Network based Model with Improved Activation Function and Optimizer for Effective Intrusion Detection and Classification," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Mar. 2021, pp. 373–378. doi: 10.1109/ICACITE51222.2021.9404584.

[19] J. O. Onah, S. M. Abdulhamid, M. Abdullahi, I. H. Hassan, and A. Al-Ghusham, "Genetic Algorithm based feature selection and Naïve Bayes for anomaly detection in fog computing environment," *Mach. Learn. with Appl.*, vol. 6, no. August, p. 100156, Dec. 2021, doi: 10.1016/j.mlwa.2021.100156.

[20] U. Sabeel, S. S. Heydari, K. Elgazzar, and K. El-Khatib, "Building an Intrusion Detection System to Detect Atypical Cyberattack Flows," *IEEE Access*, vol. 9, pp. 94352–94370, 2021, doi: 10.1109/ACCESS.2021.3093830.

[21] T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, "BAT: Deep Learning Methods on Network Intrusion Detection Using NSL-KDD Dataset," *IEEE Access*, vol. 8, pp. 29575–29585, 2020, doi: 10.1109/ACCESS.2020.2972627.

[22] M. Umair, Z. Saeed, M. Ahmad, H. Amir, B. Akmal, and N. Ahmad, "Multi-class Classification of Bi-lingual SMS using Naive Bayes Algorithm," *Proc. - 2020 23rd IEEE Int. Multi-Topic Conf. INMIC 2020*, 2020, doi: 10.1109/INMIC50486.2020.9318153.

[23] H. Liu, B. Lang, M. Liu, and H. Yan, "CNN and RNN based payload classification methods for attack detection," *Knowledge-Based Syst.*, vol. 163, pp. 332–341, Jan. 2019, doi: 10.1016/j.knosys.2018.08.036.

[24] Z. Li, A. L. G. Rios, G. Xu, and L. Trajković, "Machine learning techniques for classifying network anomalies and intrusions," *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2019-May, pp. 1–5, 2019, doi: 10.1109/ISCAS.2019.8702583.

[25] M. E. Nilă, C. Student, V. V Patriciu, and I. Bica, "Machine Learning Datasets for Cyber Security Applications," *Secur. Futur.*, vol. 3, no. 3, pp. 109–112, 2019.

[26] C. Li, J. Wang, and X. Ye, "Using a Recurrent Neural Network and Restricted Boltzmann Machines for Malicious Traffic Detection,"

NeuroQuantology, vol. 16, no. 5, May 2018, doi: 10.14704/nq.2018.16.5.1391.

[27] I. Sumaiya Thaseen and C. Aswani 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.

[28] C. Yin, Y. Zhu, J. Fei, and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017, doi: 10.1109/ACCESS.2017.2762418.

[29] C.-C. Sun, D. J. Sebastian Cardenas, A. Hahn, and C.-C. Liu, "Intrusion Detection for Cybersecurity of Smart Meters," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 612–622, Jan. 2021, doi: 10.1109/TSG.2020.3010230.

[30] M. K. Ahn, Y. H. Kim, and J. R. Lee, "Hierarchical multi-stage cyber attack scenario modeling based on G and E model for cyber risk simulation analysis," *Appl. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/app10041426.

[31] E. Mbunge, B. Muchemwa, J. Batani, and N. Mbuyisa, "A review of deep learning models to detect malware in Android applications," *Cyber Secur. Appl.*, vol. 1, p. 100014, 2023, doi: https://doi.org/10.1016/j.csa.2023.100014.

[32] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, 2019, doi: 10.1186/s42400-019-0038-7.

[33] S. Anwar *et al.*, "From Intrusion Detection to an Intrusion Response System: Fundamentals, Requirements, and Future Directions," *Algorithms*, vol. 10, no. 2, p. 39, Mar. 2017, doi: 10.3390/a10020039.

[34] L. Yang, J. Li, G. Fehringer, P. Barraclough, G. Sexton, and Y. Cao, "Intrusion detection system by fuzzy interpolation," in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Jul. 2017, pp. 1–6. doi: 10.1109/FUZZ-IEEE.2017.8015710.

[35] H. Wu *et al.*, "PD-CPS: A practical scheme for detecting covert port scans in high-speed networks," *Comput. Networks*, vol. 231, p. 109825, Jul. 2023, doi: 10.1016/j.comnet.2023.109825.

[36] R. S. S. Moorthy and N. Nathiya, “Botnet Detection Using Artificial Intelligence,” *Procedia Comput. Sci.*, vol. 218, pp. 1405–1413, 2023, doi: 10.1016/j.procs.2023.01.119.

[37] S. Das and M. J. Nene, “A survey on types of machine learning techniques in intrusion prevention systems,” in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Mar. 2017, pp. 2296–2299. doi: 10.1109/WiSPNET.2017.8300169.

[38] S. Ray, “A Quick Review of Machine Learning Algorithms,” in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Feb. 2019, pp. 35–39. doi: 10.1109/COMITCon.2019.8862451.

[39] 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.

[40] S. Loussaief and A. Abdelkrim, “Machine learning framework for image classification,” in *2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, Dec. 2016, pp. 58–61. doi: 10.1109/SETIT.2016.7939841.

[41] F. Harahap, A. Y. N. Harahap, E. Ekadiansyah, R. N. Sari, R. Adawiyah, and C. B. Harahap, “Implementation of Naïve Bayes Classification Method for Predicting Purchase,” in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, 2018, pp. 1–5. doi: 10.1109/CITSM.2018.8674324.

[42] S. Budiyanto and I. Pratama, “Classification of Network Status in Academic Information Systems using Naive Bayes Algorithm Method,” in *2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP)*, Sep. 2020, pp. 107–112. doi: 10.1109/BCWSP50066.2020.9249398.

[43] F.-J. Yang, “An Implementation of Naive Bayes Classifier,” in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, Dec. 2018, pp. 301–306. doi: 10.1109/CSCI46756.2018.00065.

[44] V. Balakrishnan and W. Kaur, “String-based Multinomial Naïve Bayes for Emotion Detection among Facebook Diabetes Community,” *Procedia Comput. Sci.*, vol. 159, pp. 30–37, 2019, doi: 10.1016/j.procs.2019.09.157.

[45] M. Artur, “Review the performance of the Bernoulli Naïve Bayes Classifier in Intrusion Detection Systems using Recursive Feature Elimination with Cross-validated selection of the best number of features,” *Procedia Comput. Sci.*, vol. 190, no. 2019, pp. 564–570, 2021, doi: 10.1016/j.procs.2021.06.066.

[46] B. Seref and E. Bostanci, “Sentiment Analysis using Naive Bayes and Complement Naive Bayes Classifier Algorithms on Hadoop Framework,” *ISMSIT 2018 - 2nd Int. Symp. Multidiscip. Stud. Innov. Technol. Proc.*, pp. 1–7, 2018, doi: 10.1109/ISMSIT.2018.8567243.

[47] D. Petschke and T. E. M. Staab, “A supervised machine learning approach using naive Gaussian Bayes classification for shape-sensitive detector pulse discrimination in positron annihilation lifetime spectroscopy (PALS),” *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 947, p. 162742, Dec. 2019, doi: 10.1016/j.nima.2019.162742.

[48] É. R. Santana, L. Lopes, and R. M. de Moraes, “Recognition of the Effect of Vocal Exercises by Fuzzy Triangular Naive Bayes, a Machine Learning Classifier: A Preliminary Analysis,” *J. Voice*, 2022, doi: 10.1016/j.jvoice.2022.10.001.

[49] Y. Wang, Y. Jia, Y. Tian, and J. Xiao, “Deep reinforcement learning with the confusion-matrix-based dynamic reward function for customer credit scoring,” *Expert Syst. Appl.*, vol. 200, no. June 2021, p. 117013, 2022, doi: 10.1016/j.eswa.2022.117013.

[50] M. K. S. Verma *et al.*, “On-Board State Estimation in Electrical Vehicles: Achieving Accuracy and Computational Efficiency Through an Electrochemical Model,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2563–2575, Mar. 2020, doi: 10.1109/TVT.2020.2966266.

[51] V. A. STAN, D. BURETEA, R. A. GHEORGHIU, and V. IORDACHE, “Simple Method to Increase Precision and Stability for Frequency Counters,” in *2020 12th International Conference on Electronics,*

Computers and Artificial Intelligence (ECAI), Jun. 2020, pp. 1–4. doi: 10.1109/ECAI50035.2020.9223216.

[52] M. Heydarian and T. E. Doyle, "MLCM : Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.

[53] D. Chen, Y. Lu, and C.-Y. Hsu, "Measurement Invariance Investigation for Performance of Deep Learning Architectures," *IEEE Access*, vol. 10, pp. 78070–78087, 2022, doi: 10.1109/ACCESS.2022.3192468.

[54] L. Koc, T. A. Mazzuchi, and S. Sarkani, "A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier," *Expert Syst. Appl.*, vol. 39, no. 18, pp. 13492–13500, 2012, doi: 10.1016/j.eswa.2012.07.009.

[55] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 2018, pp. 108–116. doi: 10.5220/0006639801080116.

[56] H. Ahmetoglu and R. Das, "A comprehensive review on detection of cyber-attacks: Data sets, methods, challenges, and future research directions," *Internet of Things*, vol. 20, p. 100615, Nov. 2022, doi: 10.1016/j.iot.2022.100615.

[57] Z. Yang *et al.*, "A systematic literature review of methods and datasets for anomaly-based network intrusion detection," *Comput. Secur.*, vol. 116, 2022, doi: 10.1016/j.cose.2022.102675.

[58] V. Kanimozhi and T. P. Jacob, "Artificial Intelligence based Network Intrusion Detection with Hyper-Parameter Optimization Tuning on the Realistic Cyber Dataset CSE-CIC-IDS2018 using Cloud Computing," in *2019 International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 33–36. doi: 10.1109/ICCSP.2019.8698029.

[59] S. Choudhary and N. Kesswani, "ScienceDirect Analysis Analysis of and UNSW-NB15 UNSW-NB15 Datasets Datasets using Deep Learning in IoT using Deep Learning in IoT," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1561–1573, 2020, doi: 10.1016/j.procs.2020.03.367.