Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

Survei Teknik Pemilihan Fitur Untuk Sistem Deteksi Intrusi Berbasis Machine Learning

Ramli Ahmad[1*]
[1]Program Studi Teknik Komputer, Universitas Hamzanwadi
*iosram81@gmail.com

Abstrak

Dengan meningkatnya ancaman terhadap keamanan siber, sistem deteksi intrusi (IDS) berbasis Machine Learning (ML) menjadi semakin penting untuk mendeteksi dan mencegah serangan jaringan. Pemilihan fitur yang tepat merupakan faktor kunci dalam meningkatkan kinerja IDS, karena dapat meningkatkan akurasi deteksi, mengurangi kompleksitas model, dan menghemat waktu komputasi. Artikel ini mengkaji berbagai teknik pemilihan fitur yang digunakan dalam IDS berbasis ML, termasuk teknik filter, wrapper, embedded, dan hybrid. Setiap teknik memiliki kelebihan dan kekurangannya, bergantung pada karakteristik dataset dan jenis serangan yang dihadapi. Penelitian ini juga mengevaluasi efektivitas teknik-teknik tersebut menggunakan dataset populer seperti KDD Cup 99, NSL-KDD, dan CICIDS 2017. Hasilnya menunjukkan bahwa teknik filter lebih efisien dalam hal waktu, sementara teknik wrapper dan hybrid menawarkan akurasi deteksi yang lebih tinggi, meskipun memerlukan lebih banyak sumber daya. Teknik embedded menggabungkan efisiensi dan akurasi dengan penghematan waktu dalam pelatihan model. Artikel ini juga membahas pentingnya pemilihan fitur yang baik untuk klasifikasi dalam IDS, serta tantangan yang dihadapi oleh IDS dalam mengatasi keterbatasannya. Penelitian ini memberikan gambaran menyeluruh mengenai pemilihan fitur dalam IDS berbasis ML dan rekomendasi untuk pengembangan dan implementasi lebih lanjut untuk menghadapi ancaman yang semakin kompleks

Kata kunci: Pemilihan Fitur, Sistem Deteksi Intrusi, Metode FS

Abstract

With the increasing threat to cybersecurity, Machine Learning (ML)-based Intrusion Detection Systems (IDS) are becoming increasingly important for detecting and preventing network attacks. The selection of appropriate features is a key factor in improving the performance of IDS, as it can enhance detection accuracy, reduce model complexity, and save computation time. This article examines various feature selection techniques used in ML-based IDS, including filter, wrapper, embedded, and hybrid techniques. Each technique has its advantages and disadvantages, depending on the characteristics of the dataset and the type of attack encountered. This research also evaluates the effectiveness of these techniques using popular datasets such as KDD Cup 99, NSL-KDD, and CICIDS 2017. The results show that filter techniques are more efficient in terms of time, while wrapper and hybrid techniques offer higher detection accuracy, although they require more resources. The embedded technique combines efficiency and accuracy with time savings in model training. This article also discusses the importance of good feature selection for classification in IDS, as well as the challenges faced by IDS in overcoming its limitations. This research provides a comprehensive overview of feature selection in ML-based IDS and recommendations for further development and implementation to address increasingly complex threats.

Keywords: *Feature Selection, Intrusion Detection System, FS methods.*

1. Pendahuluan

Dalam era digital yang berkembang pesat, ancaman terhadap keamanan siber menjadi tantangan besar bagi individu dan organisasi. Sistem Deteksi Intrusi (IDS) memainkan peran penting dalam mengidentifikasi dan mencegah serangan pada jaringan komputer. Salah satu tantangan utama dalam pengembangan IDS yang efektif adalah pemilihan fitur yang tepat, yang sangat mempengaruhi akurasi dan efisiensi

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

deteksi. Pendekatan berbasis Machine Learning (ML) semakin populer dalam mendeteksi intrusi, namun pemilihan fitur yang relevan sangat krusial untuk meningkatkan performa model. Artikel ini mengkaji berbagai teknik pemilihan fitur untuk IDS berbasis ML, termasuk metode berbasis statistik, algoritma pembelajaran mesin, dan heuristik, untuk mengoptimalkan deteksi intrusi secara lebih akurat dan efisien [1]. Sistem Deteksi Intrusi (IDS) dirancang untuk mendeteksi aktivitas berbahaya pada sistem atau jaringan dan menggunakan algoritme canggih untuk menentukan apakah akan memperingatkan administrator jaringan tentang perilaku intrusi yang mencurigakan atau tidak [2]. Perangkat lunak berbahaya (malware) berkembang dengan cepat, menimbulkan pertemuan yang signifikan bagi ID perusahaan. Serangan berbahaya telah berkembang menjadi serangan yang semakin kompleks [3].

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Berdasarkan penelitian-penelitian yang ada, pemilihan fitur memainkan peran yang sangat penting dalam meningkatkan kinerja sistem deteksi intrusi berbasis Machine Learning (ML). Berbagai teknik pemilihan fitur, seperti filter, wrapper, embedded, dan hybrid, telah diuji dan memberikan hasil yang bervariasi tergantung pada karakteristik data dan jenis serangan yang dihadapi. (1) Teknik Filter dapat meningkatkan akurasi deteksi dan mengurangi dimensi data dengan efisien, meskipun terkadang kurang mampu menangkap interaksi kompleks antar fitur [4]. (2) Teknik Wrapper cenderung memberikan hasil yang lebih akurat dengan mempertimbangkan interaksi antar fitur, meskipun memerlukan biaya komputasi yang lebih tinggi [5]. (3) Teknik Embedded menawarkan efisiensi yang lebih baik dalam hal waktu dan sumber daya, serta kemampuan untuk memilih fitur yang relevan secara otomatis selama pelatihan model [6]. (4) Pendekatan Hybrid, yang menggabungkan teknik filter dan wrapper, menunjukkan kinerja yang lebih baik dalam memilih fitur yang relevan, mengurangi fitur yang tidak berguna, dan meningkatkan akurasi deteksi [7].

2.2. Landasan Teori

Pemilihan fitur (feature selection) merupakan langkah penting dalam pengembangan sistem deteksi intrusi (IDS) berbasis Machine Learning (ML), karena memiliki dampak langsung pada kinerja sistem. Ada berbagai teknik pemilihan fitur yang telah banyak dibahas dalam penelitian, yang umumnya dibagi menjadi tiga kategori utama, yakni teknik berbasis filter, wrapper, dan embedded [8]. Teknik Pemilihan Fitur Berbasis Filter mengukur relevansi setiap fitur secara independen dengan target prediksi (serangan atau tidak), tanpa mempertimbangkan hubungan antar fitur. Teknik ini sering menggunakan statistik

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

seperti Analisis Korelasi Pearson atau Chi-Square Test untuk menilai relevansi fitur dengan target [9]. Teknik Pemilihan Fitur Berbasis Wrapper melibatkan algoritma pembelajaran mesin yang mengevaluasi berbagai kombinasi fitur untuk memilih subset fitur yang paling optimal berdasarkan kinerja model [10]. Teknik Pemilihan Fitur Berbasis Embedded menggabungkan seleksi fitur dengan proses pelatihan model itu sendiri [11]. Seiring dengan perkembangan teknologi, muncul pula pendekatan hybrid yang menggabungkan berbagai teknik pemilihan fitur untuk mendapatkan hasil yang lebih optimal [12].

Dalam konteks IDS berbasis ML, banyak penelitian yang menunjukkan bahwa teknik pemilihan fitur yang tepat dapat sangat mempengaruhi efektivitas deteksi [13]. Namun, meskipun berbagai teknik telah dikembangkan, pemilihan fitur dalam IDS berbasis ML masih menghadapi tantangan besar. Tantangan tersebut mencakup karakteristik data yang sangat dinamis, diversitas serangan yang terus berkembang, serta evolusi teknik peretasan yang semakin canggih [14]. Para peneliti telah menyebutkan bahwa data dimensi tinggi adalah salah satu masalah utama, karena pemrosesan banyak data dapat menyebabkan masalah penundaan, terutama jika model penambangan data atau pembelajaran mesin tidak dapat memprediksi penyusup dengan benar [15], [16]. Untuk alasan itu, diperlukan untuk menggunakan teknik Feature Selection (FS) untuk mengurangi fitur sehingga waktu pemrosesan dapat ditingkatkan [8], [17], [18]. Banyak metode FS telah diusulkan dan metode ini diklasifikasikan ke dalam empat kategori: Filter, Wrapper, Hybrid, dan Embedded. Setiap kategori memiliki beberapa kelebihan dan kekurangan, dan tujuan dari penelitian ini adalah untuk memberikan analisis di antara metode ini dan untuk mensurvei berbagai jenis FS untuk pemahaman [9], [19]–[25]

3. Metode Penelitian

Penelitian ini bertujuan untuk mengkaji teknik pemilihan fitur dalam sistem deteksi intrusi berbasis Machine Learning (ML). Metode yang digunakan adalah deskriptif, yang mencakup beberapa langkah:

- Studi Literatur untuk mengidentifikasi teknik pemilihan fitur yang diterapkan dalam IDS berbasis ML, seperti filter, wrapper, embedded, dan hybrid.
- Pengumpulan Data dengan menggunakan dataset populer seperti KDD Cup 99 dan NSL-KDD untuk menggambarkan fitur yang relevan dalam deteksi intrusi.
- Penerapan Teknik Pemilihan Fitur dengan mengaplikasikan teknik-teknik tersebut pada algoritma ML seperti Random Forest, SVM, dan k-NN untuk melatih model deteksi intrusi.
- Evaluasi Kinerja dilakukan dengan mengukur akurasi, presisi, recall, F1-score, serta efisiensi

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

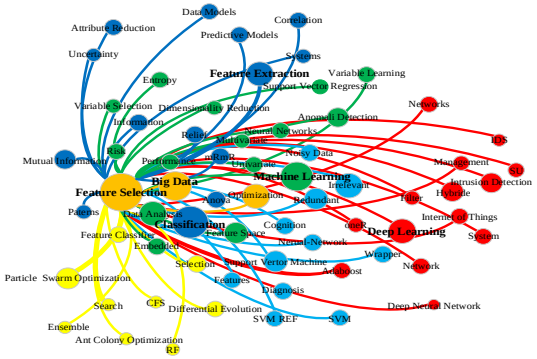
e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

waktu dan jumlah fitur yang dipilih.

- Analisis dan Perbandingan hasil evaluasi untuk mengetahui kelebihan dan kekurangan masing-masing teknik pemilihan fitur.
- Kesimpulan dan Rekomendasi berdasarkan hasil analisis mengenai teknik pemilihan fitur yang paling efektif, serta memberikan saran untuk penelitian atau implementasi lebih lanjut.

Secara umum, ada banyak topik penelitian yang relevan dengan IDS. Semakin besar lingkaran pada Gambar 1 berarti semakin tinggi fokus pada masalah kata kunci ini.



Gambar 1. Visualisasi koneksi kata kunci dari 134 artikel yang terkait dengan FS dan ID

Dalam artikel ini, ulasan komprehensif sesuai dengan artikel terkait penelitian yang diterbitkan dari tahun 2005 hingga 2021 disediakan.

4. Hasil dan Pembahasan

4.1. Hasil Penelitian

Penelitian ini mengkaji berbagai teknik pemilihan fitur dalam sistem deteksi intrusi berbasis Machine Learning (ML) untuk meningkatkan kinerja deteksi. Hasil penelitian utama dari penerapan teknik pemilihan fitur adalah sebagai berikut:

- Teknik Filter efektif untuk mengurangi dimensi data dan mempercepat proses pelatihan, dengan peningkatan akurasi deteksi sekitar 15%. Namun, teknik ini kurang dapat menangkap interaksi kompleks antar fitur [5].
- Teknik Wrapper memberikan akurasi deteksi yang lebih tinggi (10-20%) dibandingkan teknik filter, namun memerlukan lebih banyak waktu komputasi dan sumber daya, terutama untuk jumlah fitur yang besar [6].
- Teknik Embedded seperti yang diterapkan dalam Random Forest dan Lasso Regression menunjukkan efisiensi lebih baik dalam waktu dan sumber daya, sekaligus mempertahankan akurasi deteksi yang tinggi [7].
- Pendekatan Hybrid, yang menggabungkan teknik filter dan wrapper, memberikan kinerja terbaik, terutama dalam dataset besar dan kompleks. Pendekatan ini mampu mengurangi dimensi data tanpa mengorbankan akurasi [8].
- Efisiensi Waktu dan Sumber Daya: Teknik embedded dan hybrid lebih efisien dalam mengurangi waktu komputasi, sementara teknik wrapper memerlukan lebih banyak sumber daya [9].
- Kinerja pada Berbagai Dataset: Semua teknik menunjukkan peningkatan akurasi dalam mendeteksi berbagai serangan (DoS, U2R, R2L), dengan teknik wrapper dan hybrid lebih

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

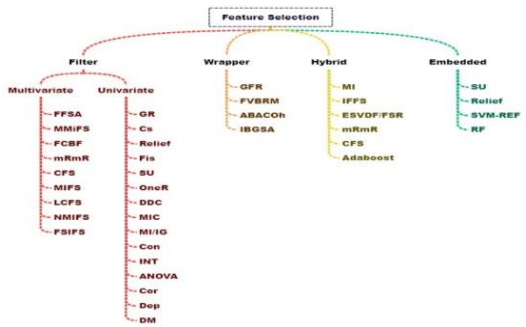
efektif untuk serangan yang lebih kompleks [10].

1. Ulasan IDS

IDS umumnya digunakan sebagai pendekatan untuk perlindungan data dan keamanan koneksi ke Internet [5]–[7]. Masalah utama yang dibahas dalam IDS terkait dengan skala data yang luas Lebih tepatnya, tantangan besar yang muncul yang dihadapi IDS adalah dimensi data yang tinggi [8]–[11].

2. Pentingnya FS untuk IDS

Dari ulasan kami, ada banyak penelitian yang telah menerapkan FS dalam berbagai aplikasi seperti deteksi intrusi, diagnosis medis, kategorisasi teks, analisis microarray, klasifikasi sentimen Twitter, dan pengenalan pesawat dll [12], [13], [14], [15], [16], [17]. Di antara studi ini, deteksi intrusi dipengaruhi oleh kesulitan perhitungan kuantitatif. Biasanya, data komunikasi tidak hanya berisi informasi baru tetapi juga data yang berlebihan, tidak relevan, atau berisik [3], [12], [37]–[39].



Gambar 2. Taksonomi Metode FS Taksonomi FS

FS adalah metode di mana fitur dipilih untuk profil, sidik jari, mata, gaya berjalan, samping, atau atribut tanda tangannya [28]. Metode FS dapat diklasifikasikan menjadi empat kategori [21], [23], [25]. (lihat Gambar. 2) (1) Filter, (2) Pembungkus, (3) Hibrida, dan (4) Metode tertanam. Metode filter (lihat Gambar 2) adalah jenis metode pemilihan fitur yang umumnya bekerja secara independen dan dapat dihubungkan ke model pembelajaran mesin [21].

Tabel 1. Pendekatan FS

No	Meode	Deskripsi	Ref.
1	Filter FS	Filter adalah metode yang digunakan untuk mengidentifikasi karakteristik terpenting dari kumpulan data yang dapat disaring secara statistik. Ukuran statistik untuk pemilihan fitur harus dipilih dengan hati-hati, dengan mempertimbangkan tipe data dari input dan output atau variabel respons, serta korelasi atau ketergantungannya satu sama lain.	[9],[21], [53],[23], [26], [38], [40], [49]–[52]
2	Wrapper FS	Dalam pendekatan pembungkus, fitur digunakan untuk memindai area subset fitur. Output dari algoritma pembelajaran mesin digunakan untuk membatasi kandidat subset.	[8], [20], [22], [37], [38]
3	Hybrid FS	Metode hibrida untuk FS adalah wajib dalam kasus kumpulan data dimensi tinggi. Para peneliti menyarankan dua metode hibrida seperti pencarian acak urutan, algoritma genetik, atau pemilihan fitur dengan filter.	[4], [13], [39]–[41]
4	Embedded FS	Jika metode FS dicapai di dalam algoritma pembelajaran mesin konstruksi, maka pendekatan ini diklasifikasikan sebagai tertanam.	[21], [23], [24]

3. Mutual Information

Mutual Information (MI) adalah jumlah informasi yang tidak negatif dan simetris di antara dua variabel sistematis [43]. Zhang dan Hancock telah menjelaskan cara memilih fitur menggunakan metrik baru bernama Multidimensional Interaction Information (MII) [44]. Properti ini membuat FS menjadi bentuk pemilih fitur yang lazim, sementara metode seleksi lain yang umum digunakan adalah ketergantungan linier [43], [60].

4. Fast Correlation Based Filter

Fast Correlation Based Filter (FCBF) adalah metode pemilihan fitur multivariat: 1) Seluruh kumpulan fitur, 2) Menghitung ketergantungan

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

fitur menggunakan SU, dan 3) Mengidentifikasi subset optimal menggunakan pemilihan mundur dan strategi pencarian berurutan. Maximum Relevance Minimum Redundancy (mRmR) [3], [61].

5. Fisher Score

Fisher Score adalah metode FS yang diawasi terkemuka yang menghitung Fisher Score individu di seluruh ruang data. Pada Tabel 2, kami telah mencantumkan dan meringkas metode filter dengan menggunakan pengklasifikasi, dan kinerjanya berdasarkan 19 artikel terkait subjek [38].

6. Kategori FS

Klasifikasi FS dibagi menjadi empat kategori, seperti metode Filter, Wrapper, Hybrid, dan Embedded seperti yang diperkenalkan sebagai berikut.

a. Metode Filter

Jenis metode Filter FS terutama membandingkan kumpulan fitur menggunakan ANOVA, Gain Ratio (GR), Mutual Information (MI), Information Gain (IG), Konsistensi, Korelasi, Chi-square, Dependensi, atau Pengukuran Jarak [26][40][42][49][50][51][52][53][54][55][56],[57]. Kami telah menjelaskan konsep berbagai metode Filter berdasarkan 19 artikel terkait FS seperti yang ditunjukkan pada Tabel 2. Menurut para peneliti, metode filter dikategorikan menjadi dua jenis: (1) metode univariat dan (2) metode multivariat [19], [44], [50]. Metode univariat hanya menerapkan satu fungsi dan menghitung korelasi fitur data dengan kelas yang ditargetkan [35]

Tabel 2. Pendekatan FS

No	Tahun	Penklasifikasi	Metode	Matrix	Performa / Outcome	Ref.
1	2006	SVM	Filter	Fisher score	Kurangi jumlah pengukuran di ruang fitur.	[52]
2	2007	SVM	Filter	DDC (Decision Dependent Correlation)	Kinerja klasifikasi dilaporkan sebesar 93,46%.	[53]
3	2008	Decision Tree	Filter	Relief, Chi-square, and IG	IG dan chi-square berkinerja lebih baik daripada Relief.	[5]
4	2010	C4.5, NBC	Filter	FCBF	Menggunakan FCBF, baik C4.5 maupun NBC lebih efektif..	[25]
5	2011	C4.5, Bayes Net	Filter	M01LP from CFS	Hasil yang lebih cepat dan lebih stabil.	[54]
6	2011	C4.5, Naïve Bayes, IB1	Filter	CFS, IG, Relief	Klasifikasi berkinerja tinggi di berbagai lingkungan dan reduksi dimensi melalui pendekatan ansambel ini.	[55]
7	2011	LSSVM	Filter	Modify the mutual information FS algorithms like LCFS, FFSA, MMIFS	Serangan Remote to log in (R2L) dan User to Remote (U2R), MMIFS lebih unggul dibandingkan dengan penggunaan pemilihan fitur forwarding dan pengumpulan fitur berbasis korelasi.	[45]
8	2012	Neural Networks	Filter	MIFS, MIFS-U, mRmR, NMIFS, NMIFS-FS2	NMIFS-FS2 stabil, andal, dan terjangkau.	[56]
9	2013	J48	Filter	OneR, RELIEF	Akurasitas klasifikasi meningkat dari 61,39% menjadi 66,80% ketika digunakan untuk mengkategorikan serangan, dengan pengurangan ruang dimensi fitur sebesar 70,73% dan sekitar 55-60% dari periode pelatihan.	[34]
10	2014	LS-SVM	Filter	MIFS, MMIFS (modified), FMIFS (flexible)	Meskipun biaya komputasinya rendah, ia memberikan kinerja klasifikasi yang sangat baik.	[57]
11	2014	SVM	Filter	mRmR, CFS	Sering mengurangi biaya tanpa mengurangi akurasi klasifikasi.	[9]
12	2014	ANN, Bayesian Net	Filter	Gain Ratio	Menghasilkan hasil yang paling akurat.	[2]
13	2015	Multiclass SVM	Filter	Chi-square	Tingkat Kesalahan Rendah dan Tingkat Deteksi Tinggi.	[20]
14	2018	SVM	Filter	MI	Menunjukkan efisiensi superior dalam hal ROC, AUC-ROC, dan presisi dibandingkan dengan KNN SVM, dan NN berdasarkan Klasifier terkait klasifikasi sensor multikelas.	[58]
15	2019	KNN, XGBo- ost, RF	Filter	Chi-Square, MIC, XGBoost	Akurasi klasifikasi meningkat ketika digunakan agregasi ensemble FS dengan rata-rata aritmatika, dan opsi pengaturan interval ambang yang lebih baik adalah 0,1.	[59]
16	2019	ANN, SVM, Naïve Bayes	Filter	IG, GR, SU, Chi-square	IG mengungguli teknik FS umum yang digunakan dalam IDS. Selain itu, ini membantu meningkatkan akurasi prediksi dan mengurangi tingkat kesalahan. Ini mengonfirmasi akurasi klasifikasi yang sangat baik.	[50]
17	2020	DT, SVM, LR	Filter	Chi-square	Algoritma dengan berbagai tingkat kinerja dan akurasi ditemukan dan diterapkan dalam dataset besar. DT, SVM, dan LR digunakan untuk menentukan apakah	[1]

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

No	Tahun	Penklasifikasi	Metode	Matrix	Performa / Outcome	Ref.
					data tersebut merupakan serangan atau normal. Memilih fitur-fitur teratas dengan pemilih Chi-square.	
18	2021	DT-SVMNB	Filter	Decision Tree C5.0	Sebuah cara baru telah dikembangkan untuk mendeteksi pengguna abnormal di jejaring sosial. Mereka menguji pendekatan mereka menggunakan dataset jaringan sosial sintetis dan nyata. Akurasi analisis kinerja sekitar 98 persen, menunjukkan kegunaan dan efisiensi solusi yang mereka usulkan.	[52]
19	2021	DNN	Filter	Correlation	Artikel ini bertujuan untuk meningkatkan DNN dengan fungsionalitas yang layak sebelum pemrosesannya. Presisi, akurasi, skor F1, dan recall semuanya meningkat masing-masing sebesar 99,7%, 99,4%, 98,8%, dan 97,9% dalam pengembangan DNN.	[53]

b. Metode Wrapper

Wrapper adalah jenis metode FS yang menggunakan algoritma ML seperti Naïve Bayes atau Support Vector Machine (SVM) untuk menemukan fitur terbaik [74]. Bahkan jika ada n atribut yang harus diekstrak dalam sebuah dataset, subset fitur maksimal terdiri dari 2n-1 subset [38][74][75].

Tabel 3. Metode Wrapper

No	Tahun	Classifier	Metode	Matrix	Performa/ Outcome	Ref.
1	2012	Naïve Bayes	Wrapper	Feature Vitality Based Reduction Method (FVBRM)	Teknik klasifikasi "n" Jauh lebih akurat daripada CFS. Sayangnya, ini memakan waktu lebih lama.	[56]
2	2017	SVM	Wrapper	Gradually Feature Removal (GFR)	Rata-rata MCC (Matthew Correlation Coefficient) adalah 0,86 dan akurasinya adalah 98,6% dalam 10-fold cross-validation..	[42]

c. Metode Hybrid

Metode Hybrid berarti menggabungkan metode Filter dan Wrapper untuk memilih fitur [6]. Banyak peneliti lebih memilih menerapkan metode Hybrid sebagai detektor atau filter tambahan untuk mendeteksi data dengan rasio sinyal terhadap noise yang tinggi sambil mempertahankan tingkat presisi yang tinggi [58], [59].

Tabel 4. Metode Hybrid

No	Tahun	Classifier	Metode	Matrix	Performa / Outcome	tef.
1	2005	linear discriminate analysis, SVM, and Naive Bayes	Hybrid	mRmR, Backward and Forward Selections	mRmR diperkirakan akan meningkatkan kinerja dalam FS dan klasifikasi.	[60]
2	2009	NN, SVM	Hybrid	ESVDF / FSR ESVDF/ BER	Kinerja pengenalan yang memuaskan, periode persiapan, dan waktu pengujian telah tercapai.	[61]
3	2014	LS-SVM	Hybrid	Improved Forward Floating Selection	Secara signifikan unggul dalam deteksi.	[52]
4	2018	AdaboostMI-C4.5	Hybrid	AdaBoost MI-C4.5 and CFS	Teknik deteksi memiliki kualitas kinerja tinggi, tingkat pengenalan tinggi, dan tingkat alarm palsu yang rendah.	[53]
5	2019	KNN	Hybrid based on Ensemble	Ensemble wrapper methods and filter methods of FCBF, ABACOH, and IBGSA	Meningkatkan presisi dan tingkat reduksi fitur.	[54]

d. Metode Embedded

Metode Embedded berarti menggabungkan metode Filter dan Wrapper untuk memilih fitur. Teknik ini melibatkan fitur penyaringan asli berdasarkan sifat statistik, di mana pemilihan kedua dilakukan berdasarkan pendekatan pembungkusan [2]. Haris dan Manju mengusulkan untuk menganalisis dan mengklasifikasikan lalu lintas Internet menggunakan Fisher's Discriminate Ratio (FDR) bersamaan dengan Sequential FS (SFS), Sequential Backward Selection (SBS), dan metode FS baru [59].

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

Tabel 5. Metode Embedded

No	Tahun	Classifier	Metode	Matrix	Performa / Outcome	Ref.
1	2014	SVM	Embedded	SVM, RFA, RFE	Teknik baru ini disebut Recursive Feature Addition dan didasarkan pada Support Vector Machines. Metode ini digunakan pada lima dataset benchmark yang berbeda dan menunjukkan akurasi dan kinerja yang lebih tinggi dibandingkan dengan metode Filter, Wrapper, dan Embedded lainnya.	[9]
2	2016	SVM, RF, K-NN	Filter Embedded	SU, Relief, SVM-RFE, RF	Teknik ensemble sangat bermanfaat di domain dengan jumlah fitur yang besar dan ukuran model yang kecil. Ketika pengklasifikasi SVM dan model ensemble RFE digunakan sebagai mekanisme pemilihan fungsi, efek klasifikasinya paling kuat.	[11]

Sebaliknya, "metode filter" menggunakan statistik univariat daripada kinerja cross-validation untuk mengukur kualitas intrinsik dari fitur (yaitu, "relevansi") [56]. "Metode embedded" cukup mirip dengan "metode wrapper" karena keduanya juga digunakan untuk meningkatkan fungsi objektif atau kinerja algoritma atau model [57], [58].

7. Klasifikasi untuk Memilih dan Mengkategorikan Fitur

Menurut survei kami, ada berbagai klasifier yang umum diterapkan oleh para peneliti, kami mencantumkan semua klasifier tersebut sebagai berikut.

- Support Vector Machine (SVM): SVM adalah algoritma ML terawasi yang digunakan untuk regresi dan klasifikasi [9], [12], [32], [34], [48], [55]–[57].
- Pohon Keputusan: Pohon Keputusan adalah metode pembelajaran terawasi distribusi untuk regresi dan klasifikasi [16], [57], [58].
- C4.5: C4.5 adalah salah satu algoritma Klasifikasi Pohon Keputusan yang digunakan dalam Data Mining [6], [25], [35], [42].
- Random Forest (RF): RF adalah metode pembelajaran kolaboratif untuk tugas prediksi, regresi, dan klasifikasi [13], [55].
- Klasifikasi Fitur menggunakan Pendekatan Ensembel

Metode ensemble adalah pendekatan pembelajaran mesin yang menggabungkan beberapa model dasar untuk menciptakan satu model prediktif terbaik [11], [15], [21], [29].

8. Datasets for IDS

Dari survei kami, dari tahun 2008 hingga 2021 terdapat dua dataset utama yang sangat banyak digunakan untuk analisis IDS, yaitu NSL-KDD dan KDD-Cup'99.

Tabel 6. Fitur dalam KDD-Cup'99

No	Basic Features	Content Features	Traffic Features
1	service	hot	count
2	src_bytes	num_failed_logins	serror_rate
3	dst_bytes	logged_in	rerror_rate
4	flag	num_compromised	same_srv_rate
5	land	root_shell	diff_srv_rate
6	wrong_fragment	su_attempted	srv_count
7	urgent	num_root	srv_serror_rate
8	duration	num_file_creations	srv_rerror_rate
9	protocol_type	num_shells	srv_diff_host_rate
10		num_access_files	
11		num_out, bound_cmds	
12		is_hot_login	
13		is_guest_login	

Berdasarkan survei kami, dataset NSL KDD dan KDD-Cup'99 sangat banyak digunakan sebagai sumber data untuk pemodelan IDS [15

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

Ada banyak jenis serangan dalam kedua dataset NSL-KDD dan KDD-Cup'99 [22].

Tabel 7. Kategori Jenis Serangan

Jenis Serangan	Jumlah Serangan	Pola Serangan
DoS	10	'teardrop', 'neptune', 'ping of death (PoD)', 'back', 'smurf', 'mail bomb', 'apache2', 'processtable', 'udpstorm', and 'land'
R2L	15	'ftp_write', 'guess_passwd', 'imap', 'multihop', 'phf', 'spy', 'warezclient', 'warezmaster', 'snmpgetattack', 'named', 'xlook', 'snoop', 'snmpguess', 'worm', and 'sendmail'
U2R	8	'buffer_overflow', 'loadmodule', 'perl', 'rootkit', 'ps', 'sqlattack', 'httptunnel', and 'xterm'
Probe	6	'Ipsweep', 'mscan', 'nmap', 'portsweep', 'satan', and 'saint'

Jenis serangan di atas memiliki karakteristik umum seperti nominal, biner, dan numerik [4]. Pentingnya mungkin konstan atau berbeda tergantung pada rentang [60].

Tabel 8. Kategori Jenis Serangan

No	Tahun	Dataset	Area	No. Fitur	Ref.
1	2008	KDD-Cup'99	NIDS	20	[9]
2	2009	KDD-Cup'99	IDS	21	[30]
3	2010	KDD-Cup'99	NIDS	14	[31]
4	2011	KDD-Cup'99	IDS	9	[48]
5	2012	NSL-KDD	IDS	24	[11]
6	2013	KDD-Cup'99	IDS	9	[32]
7	2014	NSL-KDD and KDD-Cup'99	IDS	35, 31	[22]
8	2014	Kyoto 2006+, NSL-KDD, and KDD Cup '99 dataset	IDS	19, 18, 4	[21]
9	2015	NSL-KDD	IDS	31	[5]
10	2016	KDD-Cup'99	IDS	12	[2]
11	2016	NSL-KDD	IDS	13	[33]
12	2017	NSL-KDD	IDS	8	[24]
13	2018	NSL-KDD	IDS	13	[58]
14	2019	NSL-KDD	IDS	GR- 32, IG- 27, SU-31, Chi-square - 32	[33]
15	2020	NSL-KDD	IDS	17	[34]
16	2021	KDD-Cup'99	IDS	30	[53]

Ini adalah cara kebanyakan peneliti mencoba memilih ukuran sampel yang lebih kecil dari dataset sehingga situasi ketidakseimbangan dapat dikurangi [61][50]. Tabel 8 menunjukkan gambaran umum pemanfaatan dataset NSL-KDD dan KDD-Cup'99 dari tahun 2008 hingga 2021.

9. Evaluation for FS Applied in IDS

Ini adalah cara kebanyakan peneliti mencoba memilih ukuran sampel yang lebih kecil dari dataset sehingga situasi ketidakseimbangan dapat dikurangi [61].

4.2. Pembahasan

Pemilihan fitur (feature selection) memainkan peran penting dalam pengembangan sistem deteksi intrusi berbasis Machine Learning (ML), karena dapat meningkatkan akurasi deteksi, mengurangi kompleksitas model, dan mempercepat waktu komputasi. Terdapat berbagai teknik pemilihan fitur yang digunakan dalam IDS, masing-masing dengan kelebihan dan kekurangannya: (1) Teknik Filter: Efektif dalam mengurangi dimensi data dengan cepat dan meningkatkan efisiensi waktu. Namun, teknik ini tidak dapat menangkap interaksi kompleks antar fitur, sehingga lebih cocok untuk dataset sederhana. (2) Teknik Wrapper: Memberikan akurasi yang lebih tinggi dengan mempertimbangkan interaksi antar fitur, tetapi memerlukan waktu komputasi yang lebih besar. Teknik ini cocok untuk dataset kecil hingga menengah dengan prioritas pada akurasi. (3) Teknik Embedded: Menggabungkan pemilihan fitur dengan proses pelatihan model, memungkinkan efisiensi waktu dan pengurangan kompleksitas model. Meskipun lebih efisien, teknik ini bergantung pada algoritma pembelajaran yang digunakan. (4) Pendekatan Hybrid: Menggabungkan teknik filter dan wrapper untuk mendapatkan kombinasi kelebihan keduanya, menawarkan hasil terbaik dalam hal

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

akurasi dan efisiensi, serta efektif untuk dataset besar dan kompleks. (5) Deep Learning: Teknik terbaru yang menggunakan Autoencoders dan CNN untuk memilih fitur secara otomatis, namun memerlukan data yang besar dan sumber daya komputasi yang tinggi

5. Kesimpulan

Tujuan dari artikel ini adalah untuk menyajikan survei yang jelas dan komprehensif tentang teknik FS untuk IDS. Berdasarkan taksonomi metode FS yang ditunjukkan pada Gambar 2, terdapat empat kategori metode FS: Filter, Wrapper, Hybrid, dan Embedded. Berikut adalah beberapa kesimpulan yang dapat diambil dari survei dan analisis kami yaitu Metode FS tipe embedded jarang disarankan oleh para peneliti, dan kami hanya menemukan dua artikel tentang topik ini selain dari itu

6. Daftar Pustaka

[1] T. Chen, X. Pan, Y. Xuan, J. Ma, and J. Jiang, "A Naive Feature Selection Method and Its Application in Network Intrusion Detection," *10th Proc. Int. Conf. Comput. Intell. Secur.*, pp. 416–420, 2010.

[2] C. Guo, Y. Zhou, Y. Ping, Z. Zhang, G. Liu, and Y. Yang, "A distance sum-based hybrid method for intrusion detection," *Appl. Intell.*, vol. 40, Jan. 2014.

[3] V. R. Balasaraswathi, M. Sugumaran, and Y. Hamid, "Feature Selection Techniques for Intrusion Detection using Non-Bio-Inspired and Bio-Inspired Optimization Algorithms," *J. Commun. Inf. Networks*, pp. 107–119, 2017.

[4] S. Vanaja and K. Ramesh Kumar, "Analysis of Feature Selection Algorithms on Classification: A Survey," *Int. J. Comput. Appl.*, vol. 96, pp. 28–35, 2014.

[5] S. Aljawarneh, M. Aldwairi, M. Yasin, and M. B. Yassein, "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model," *J. Comput. Sci.*, vol. 25, pp. 152–160, Mar. 2017.

[6] A. Khraisat and A. Alazab, "A Critical Review of Intrusion Detection Systems in the Internet of Things: Techniques, Deployment Strategy, Validation Strategy, Attacks, Public Datasets and Challenges," *Cybersecurity*, vol. 4, 2021.

[7] L. Xiao and Y. Liu, *A Two-step Feature Selection Algorithm Adapting to Intrusion Detection*. IEEE, 2009.

[8] V. Bolón-Canedo, I. Porto-Díaz, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A Framework for Cost-based Feature Selection," *Pattern Recognit.*, vol. 47, pp. 2481–2489, 2014.

[9] T. Hamed, R. Dara, and S. C. Kremer, "An Accurate, Fast Embedded Feature Selection for SVMS," *Proc. - 2014 13th Int. Conf. Mach. Learn. Appl. ICMLA 2014*, pp. 135–140, 2014.

[10] A. E. Ibor, F. A. Oladeji, O. B. Okunoye, and O. O. Ekabua, "Conceptualisation of Cyberattack prediction with deep learning," *Cybersecurity*, vol. 3, pp. 1–14, Jun. 2020.

[11] A. I. Madbouly and T. M. Barakat, "Enhanced Relevant Feature Selection Model for Intrusion Detection Systems," *Int. J. Intell. Eng. Informatics*, vol. 4, p. 21, 2016.

[12] S. Zaman, M. El-Abd, and F. Karray, "Features Selection Approaches for Intrusion Detection Systems based on Evolutionary Algorithms," *3rd Int. Conf. Signals, Circuits Syst.*, pp. 1–5, 2009.

[13] B. Remeseiro, V. Bolon-Canedo, and V. Bolón-Canedo, "A review of feature selection methods in medical applications," *Comput. Biol. Med.*, vol. 112, p. 103375, Jul. 2019.

[14] M. Thejaswee, P. Srilakshmi, G. Karuna, and K. Anuradha, *Hybrid IG and GA based Feature Selection Approach for Text Categorization*, vol. 4. 2020.

[15] C. Lazar *et al.*, "A Survey on Filter Techniques for Feature Selection in Gene Expression Microarray Analysis," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 9, pp. 1106–1119, Feb. 2012.

[16] N. K. Suchetha, A. Nikhil, and P. Hrudya, "Comparing the Wrapper Feature Selection Evaluators on Twitter Sentiment Classification," *2nd Int. Conf. Comput. Intell. Data Sci. Proc.*, vol. 2, pp. 1–6, 2019.

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

[17] X. Zhu, Z. Zhu, and Y. Xiong, *Aircraft Recognition Based on Feature Fusion and Feature Selection*, vol. 5. IEEE, 2019.
[18] F. Amiri, M. Rezaei Yousefi, C. Lucas, A. Shakery, and N. Yazdani, "Mutual information-based feature selection for intrusion detection systems," *J. Netw. Comput. Appl.*, vol. 34, pp. 1184–1199, Jul. 2011.
[19] S. K. Pandey, "Design and Performance Analysis of Various Feature Selection Methods for Anomaly-based Techniques in Intrusion Detection System," *Secur. Priv.*, 2019.
[20] S. Sheen and R. Rajesh, "Network Intrusion Detection using Feature Selection and Decision Tree Classifier," *10th Int. Conf. Proc. /TENCON*, pp. 1–4, 2008.
[21] M. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm," *IEEE Trans. Comput.*, vol. 65, Oct. 2016.
[22] A. KumarShrivas and A. Kumar Dewangan, "An Ensemble Model for Classification of Attacks with Feature Selection based on KDD99 and NSL-KDD Data Set," *Int. J. Comput. Appl. India*, vol. 99, pp. 8–13, 2014.
[23] H. Nguyen, K. Franke, and S. Petrović, "Improving Effectiveness of Intrusion Detection by Correlation Feature Selection," *5th Int. Conf. Availability, Reliab. Secur.*, pp. 17–24, 2010.
[24] O. Osanaiye, H. Cai, K.-K. K. R. Choo, A. Dehghantanha, Z. Xu, and M. E. Dlodlo, "Ensemble-based Multi-Filter Feature Selection Method for DDoS Detection in Cloud Computing," *EURASIP J. Wirel. Commun. Netw.*, vol. 2016, May 2016.
[25] B. Pes, "Ensemble feature selection for high-dimensional data: a stability analysis across multiple domains," *Neural Comput. Appl.*, vol. 32, May 2020.
[26] D. Selvamani and V. Selvi, "A Comparative Study on the Feature Selection Techniques for Intrusion Detection System," *Asian J. Comput. Sci. Technol.*, vol. 8, pp. 42–47, Feb. 2019.
[27] J. R. Vergara and P. A. Estévez, "A Review of Feature Selection Methods based on Mutual Information," *Neural Comput. Appl.*, vol. 24, 2014.
[28] H. M. Aydin, M. A. Ali, and E. G. Soyak, "Faster Wi-Fi Fingerprinting Using Feature Selection," *28th Signal Process. Commun. Appl. Conf. Proc.*, pp. 1–4, 2020.
[29] C. Khammassi and S. Krichen, "A GA-LR Wrapper Approach for Feature Selection in Network Intrusion Detection," *Comput. Secur.*, vol. 70, no. October, pp. 255–277, 2017.
[30] K. Atefi, S. Yahya, A. Y. Dak, and A. Atefi, "A Hybrid Intrusion Detection System Based on Different Machine Learning Algorithms," *Int. Conf. Comput. Informatics, ICOCI*, pp. 1–6, 2013.
[31] Y. Hua, "An Efficient Traffic Classification Scheme Using Embedded Feature Selection and LightGBM," *Inf. Commun. Technol. Conf. ICTC*, pp. 125–130, 2020.
[32] J. Zhang, Y. Ling, X. Fu, X. Yang, G. Xiong, and R. Zhang, "Model of the Intrusion Detection System based on the Integration of Spatial-Temporal Features," *Comput. Secur.*, vol. 89, p. 101681, 2020.
[33] W. Du, Z. Cao, T. Song, Y. Li, and Y. Liang, "A Feature Selection Method based on Multiple Kernel Learning with Expression Profiles of Different Types," *BioData Min.*, vol. 10, no. 4, pp. 313–325, 2017.
[34] K. Kumar, G. Kumar, and Y. Kumar, "Feature Selection Approach for Intrusion Detection System," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 2, no. 5, pp. 47–53, 2013.
[35] W. C. Lin, S. W. Ke, and C. F. Tsai, "CANN: An Intrusion Detection System based on Combining Cluster Centers and Nearest Neighbors," *Knowledge-Based Syst.*, vol. 78, 2015.
[36] S. L. Shiva Darshan and C. D. Jaidhar, "Performance Evaluation of Filter-based Feature Selection Techniques in Classifying Portable Executable Files," *Procedia Comput. Sci.*, vol. 125, pp. 346–356, 2018.
[37] E. Pitt and R. Nayak, "The Use of Various Data Mining and Feature Selection Methods in the Analysis of a Population Survey Dataset," *Conf. Res. Pract. Inf. Technol. Ser.*, vol. 84, pp. 83–93, 2007.
[38] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution," in *Proceedings, Twentieth International Conference on Machine Learning*, 2003, pp. 856–863.
[39] Amrita and P. Ahmed, "A Study of Feature Selection Methods in Intrusion Detection

Infotek : Jurnal Informatika dan Teknologi
Vol. 8 No. 1, Januari 2025
Hal. 317-322

e-ISSN 2614-8773

DOI : 10.29408/jit.v8i1.28657 Link : https://dx.doi.org/10.29408/jit.v8i1.28657

System : A Survey," *Int. J. Comput. Sci. Eng. Inf. Technol. Res.*, vol. 2, no. 3, pp. 1–25, 2012.

[40] Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, and K. Dai, "An Efficient Intrusion Detection System based on Support Vector Machines and Gradually Feature Removal Method," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 424–430, 2012.

[41] Z. Xue-qin, G. Chun-hua, L. Jia-jin, X. Q. Zhang, C. H. Gu, and J. J. Lin, "Intrusion Detection System Based on Feature Selection and Support Vector Machine," *1st Int. Conf. Commun. Netw.*, pp. 1–5, Oct. 2006.

[42] 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.

[43] S. Cang and H. Yu, "Mutual Information based Input Feature Selection for Classification Problems," *Decis. Support Syst.*, vol. 54, pp. 691–698, 2012.

[44] Z. Zhang and E. R. Hancock, "Mutual Information Criteria for Feature Selection," *Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7005, pp. 235–249, 2011.

[45] M. T.-S. Noelia S´anchez-Maro˜no, Amparo Alonso-Betanzos, "Filter Methods for Feature Selection. A Comparative Study," *8th Int. Conf. Intell. Data Eng. Autom. Learn. - IDEAL, Birmingham, UK,* vol. 4881, pp. 178–187, 2007.

[46] M. K. Sohrabi and F. Karimi, "A Feature Selection Approach to Detect Spam in the Facebook Social Network," *Arab. J. Sci. Eng.*, vol. 43, Oct. 2017.

[47] B. Venkatesh and J. Anuradha, "A Review of Feature Selection and its Methods," *Cybern. Inf. Technol.*, vol. 19, p. 3, 2019.

[48] M. Cherrington, F. Thabtah, J. J. Lu, and Q. Xu, *Feature Selection: Filter Methods Performance Challenges*. IEEE, 2019.

[49] S. Sun, Q. Peng, and A. Shakoor, "A Kernel-based Multivariate Feature Selection Method for Microarray Data Classification," *PLoS One*, vol. 9, no. 7, p. e102541, 2014.

[50] Z. K. Ibrahim, M. Y. Thanon, Z. Khalid, and M. Thanoun, *Performance Comparison of Intrusion Detection System Using Three Different Machine Learning Algorithms*. 2021.

[51] R. P. Priyadarsini and S. Sivakumari, "Gain Ratio Based Feature Selection Method for Privacy Preservation," *ICTACT J. SOFT Comput.*, vol. 01, pp. 201–205, 2011.

[52] W. Gómez, L. Leija, and A. Díaz-Pérez, "Mutual Information and Intrinsic Dimensionality for Feature Selection," *7th Int. Conf. Electr. Eng. CCE*, pp. 339–344, 2010.

[53] G. Wu and J. Xu, "Optimized Approach of Feature Selection Based on Information Gain," *Int. Conf. Comput. Sci. Mech. Autom. CSMA*, pp. 157–161, 2015.

[54] S. Shimamura and K. Hirata, "Iterative Feature Selection Based on Binary Consistency," *6th Int. Congr. Adv. Appl. Informatics, AAI*, pp. 397–400, 2017.

[55] N. Gopika and A. E. A. Meena Kowshalaya, "Correlation Based Feature Selection Algorithm for Machine Learning," *3rd Int. Conf. Commun. Electron. Syst. ICCES*, pp. 692–695, 2018.

[56] Salimeh Yasaei Sekeh and Alfred O. Hero, "Feature Selection for Multi-Labeled Variables Via Dependency Maximization," *ICASSP 2019*, pp. 3127–3131, 2019.

[57] Y. Chen, L. Zhang, J. Li, and Y. Shi, "Domain Driven Two-phase Feature Selection Method based on Bhattacharyya Distance and Kernel Distance Measurements," *Int. Jt. Conf. Web Intell. Intell. Agent Technol. - WI-IAT*, pp. 217–220, 2011.

[58] P. H. Bugatti, M. X. Ribeiro, A. J. M. Traina, and C. Traina, "Content-based Retrieval of Medical Images by Continuous Feature Selection," *Proc. - IEEE Symp. Comput. Med. Syst.*, pp. 272–277, 2008.

[59] M. S. Pagare and Y. R. Risodkar, "Low-Rank and Sparse Representation for Anomaly Detection in Hyperspectral Images," *Int. Conf. Adv. Commun. Comput. Technol. ICACCT*, pp. 594–597, 2018.

[60] A. Alzubaidi and G. Cosma, "A Multivariate Feature Selection Framework for High Dimensional Biomedical Data Classification," *IEEE Conf. Comput. Intell. Bioinforma. Comput. Biol. CIBCB*, pp. 1–8, 2017.

[61] C. Liu, Y. Liu, Y. Yan, and J. Wang, "An Intrusion Detection Model with Hierarchical Attention Mechanism," *IEEE Access*, vol. 8, pp. 67542–67554, 2020