

**OPTIMALISASI *MULTI-CLASSIFICATION*
SERANGAN *CYBER* MENGGUNAKAN METODE
K-NEAREST NEIGHBOR**

SKRIPSI

**Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**



OLEH :

ZULI YANTI

09011182025014

**JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2024

HALAMAN PENGESAHAN

**OPTIMALISASI *MULTI-CLASSIFICATION*
SERANGAN *CYBER* MENGGUNAKAN METODE
K-NEAREST NEIGHBOR**

SKRIPSI

**Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**

Oleh :

ZULI YANTI

09011182025014

Indralaya, 9 April 2024

Mengetahui,

Ketua Jurusan Sistem Komputer

Pembimbing Tugas Akhir



Dr. Dr. Sukemi, M.T.

NIP. 196612032006041001

Ahmad Hervanto, S.Kom., M.T.

NIP. 198701222015041002

AUTHENTICATION PAGE

MULTI-CLASSIFICATION OPTIMIZATION

CYBER ATTACKS USING METHODS

K-NEAREST NEIGHBOR

SKRIPSI

Submitted To Complete One Of The Requirments For
Obtaining A Bachelor's Degree In Computer Science

By :

ZULI YANTI

09011182025014

Indralaya, 9 April 2024

Acknowledge,

Head Of Computer System
Department



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

Final Project Advisor

A handwritten signature in black ink, appearing to read 'A. Heryanto', is written over the text.

Ahmad Heryanto, S.Kom., M.T.

NIP. 198701222015041002

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Jumat

Tanggal : 22 Maret 2024

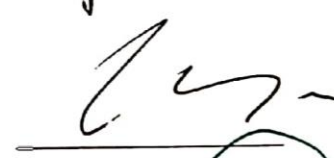
Tim Penguji :

1. Ketua : Dr. Firdaus, M.Kom

2. Sekretaris : Iman Saladin B. Azhar, M.MSI

3. Penguji : Huda Ubaya, M.T

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



Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang Bertanda Tangan Di Bawah Ini :

Nama : Zuli Yanti
NIM : 09011182025014
Judul : OPTIMALISASI *MULTI-CLASSIFICATION* SERANGAN
CYBER MENGGUNAKAN METODE K-NEAREST
NEIGHBOR

Hasil Pengecekkam Software iThenticate/Turnitin : 5%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya saya 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, 4 April 2024

Penulis,



Zuli Yanti

NIM.09011182025014

KATA PENGANTAR

Assalamualaikum Warahmatulahi Wabarakatuh

Puji syukur Alhamdulillah penulis panjatkan atas kehadiran Allah SWT, karena berkat Rahmat dan Karunia-Nya lah sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul **“Optimalisasi Multi-Classification Serangan Cyber Menggunakan Metode K-Nearest Neighbor”**. Shalawat beriringkan salam senantiasa tercurahkan kepada Nabi Muhammad Shallallaahu ‘Alaihi Wasallam yang telah membawa kedamaian dan rahmat untuk semesta alam serta menjadi suri tauladan bagi umatnya.

Pada kesempatan ini penulis menyampaikan banyak ucapan terimakasih kepada semua pihak yang telah membantu, memberikan motivasi, kemudahan, pengarahan, bimbingan, dorongan semangat, kritik dan saran selama proses penyusunan Tugas Akhir ini. Oleh karena itu, kesempatan ini penulis bersyukur dan mengucapkan banyak terima kasih kepada :

1. Allah Subhanahu Wata’ala yang telah memberikan berkah serta nikmat Kesehatan yang berlimpah.
2. Untuk diriku yang sudah berjuang dan melewati masalah – masalah yang sulit.
3. Kedua Orang Tua dan Keluarga yang selalu mendoakan serta memberikan motivasi dan semangat.
4. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Pengganti Antar Waktu Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Julian Supardi, S.Pd., M.T. selaku Wakil Dekan Bidang Akademik di Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Bapak Dr. Ir. H. Sukemi., M.T. selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Bapak Dr. Ahmad Zarkasi, M.T. selaku Dosen Pembimbing Akademik.

8. Bapak Ahmad Heryanto,S.Kom.,M.T. Selaku Dosen Pembimbing Tugas Akhir.
9. Bapak Yopi Syaputra selaku Admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal-hal administrasi.
10. Para team MC yang selalu membantu dan menemani penulis dalam susah maupun senang selama penelitian ini.
11. Teman-teman seperjuangan Jurusan Sistem Komputer Angkatan 2020, yang sudah menjadi support system saya.
12. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan semangat serta do'a.
13. Almamater.

Dalam penulisan Laporan Tugas Akhir ini penulis menyadari bahwa pada laporan ini masih jauh dari kesempumaan, oleh karena itu penulis sangat mengharapkan kritik dan saran dari semua pihak berkenan agar menjadi bahan evaluasi dan laporan ini menjadi lebih baik lagi dikemudian hari.

Akhir kata penulis ucapkan dan berharap semoga Laporan Tugas Akhir ini bermanfaat serta dapat memberikan pengetahuan dan wawasan bagi semua pihak yang membutuhkannya. Khususnya bagi mahasiswa/i Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya. Akhir kata saya ucapkan,

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Indralaya, 4 April 2024

Penulis,



Zuli Yanti

NIM. 09011182025014

**OPTIMALISASI MULTI-CLASSIFICATION SERANGAN CYBER
MENGUNAKAN METODE K-NEAREST NEIGHBOR**

ZULI YANTI

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
Email : zuliyanti529@gmail.com

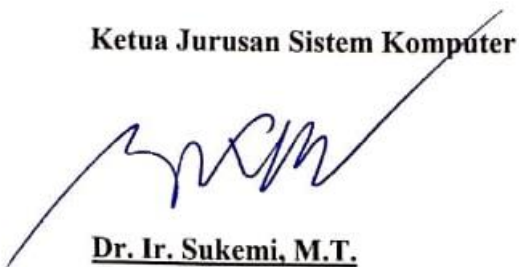
ABSTRAK

Serangan siber adalah bentuk ancaman yang bertujuan mencuri, merusak, atau mengubah data penting dalam sistem komputer atau jaringan. Jenis serangan ini meliputi peretasan, phishing, dan malware. Mendeteksi serangan siber menjadi krusial untuk menjaga keamanan sistem informasi dari ancaman yang semakin meningkat. Salah satu alat yang digunakan untuk deteksi ancaman tersebut adalah Sistem Deteksi Intrusi (IDS), yang memantau peristiwa dalam sistem dan bertindak jika ada serangan atau perilaku mencurigakan. Dalam upaya meningkatkan kinerja IDS, penelitian telah mengeksplorasi penggunaan kecerdasan buatan (AI), terutama teknik Machine Learning (ML). Dalam penelitian ini, akan dibahas penerapan metode K-NN (K-Nearest Neighbors), suatu teknik nonparametrik untuk mengukur jarak antara data baru dengan data yang sudah diklasifikasikan sebelumnya menggunakan Euclidean distance. Untuk menguji keefektifan metode K-NN, berbagai dataset digunakan, termasuk UNSW-NB15, NSL-KDD, ISCX2012, dan CIC-IDS-2018. Hasil penelitian menunjukkan bahwa model berhasil mencapai tingkat akurasi yang tinggi pada masing-masing dataset selama proses pelatihan, yakni 88.00% pada dataset UNSW-NB15, 99.99% pada dataset NLS-KDD, 99.96% pada dataset ISCX2012, dan 92.00% pada dataset CIC-IDS-2018.

Kata Kunci : KNN, Serangan Siber, UNSW-NB15, NSL-KDD, ISCX2012, CIC-IDS-2018

Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

Pembimbing Tugas Akhir



Ahmad Hervanto, S.Kom, M.T.

NIP. 198701222015041002

**MULTI-CLASSIFICATION OPTIMIZATION CYBER ATTACKS
USING METHODS K-NEAREST NEIGHBOR**

ZULI YANTI

Department of Computer Systems, Faculty of Computer Science, Sriwijaya
University

Email : zuliyanti529@gmail.com

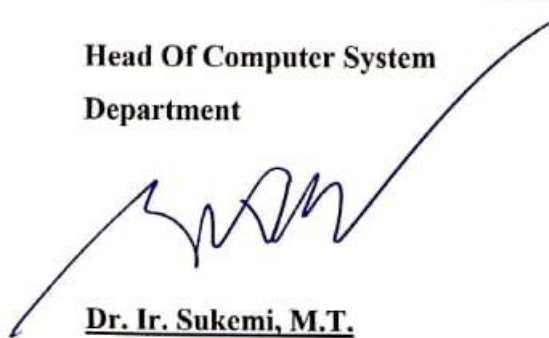
ABSTRACT

Cyber attacks are a form of threat aimed at stealing, damaging, or altering important data within computer systems or networks. These attacks include hacking, phishing, and malware. Detecting cyber attacks is crucial to maintaining the security of information systems from increasing threats. One tool used for threat detection is Intrusion Detection Systems (IDS), which monitor system events and take action if there are suspicious activities or attacks. In efforts to enhance IDS performance, research has explored the use of Artificial Intelligence (AI), particularly Machine Learning (ML) techniques. This study focuses on implementing the K-Nearest Neighbors (K-NN) method, a non-parametric technique for measuring the distance between new and previously classified data using Euclidean distance. To test the effectiveness of the K-NN method, various datasets are utilized, including UNSW-NB15, NSL-KDD, ISCX2012, and CIC-IDS-2018. The research findings indicate that the model achieves high accuracy rates on each dataset during the training process, namely 88.00% on the UNSW-NB15 dataset, 99.99% on the NSL-KDD dataset, 99.96% on the ISCX2012 dataset, and 92.00% on the CIC-IDS-2018 dataset.

Keywords: KNN, Cyber Attacks, UNSW-NB15, NSL-KDD, ISCX2012, CIC-IDS-2018.

Mengetahui,

**Head Of Computer System
Department**



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

Final Project Advisor



Ahmad Hervanto, S.Kom, M.T.

NIP. 198701222015041002

DAFTAR ISI

HALAMAN PENGESAHAN	i
AUTHENTICATION PAGE	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERNYATAAN	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xviii
DAFTAR LAMPIRAN	xxii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Tujuan	3
1.3. Manfaat	4
1.4. Perumusan Masalah	4
1.5. Batasan Masalah.....	4
1.6. Metodologi Penelitian	5
1.7. Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1. Penelitian Terdahulu	8
2.2. Timeline Penelitian Terdahulu.....	16
2.3. Ringkasan Hasil Kajian Literatur.....	18

2.4.	Serangan Siber	27
2.5.	Jenis – Jenis Serangan Siber	27
2.5.1.	Fuzzers attack	27
2.5.2.	Exploits attack	28
2.5.3.	Analysis attack.....	28
2.5.4.	Backdoors attack.....	28
2.5.5.	Generic attack	29
2.5.6.	Reconnaissance attack	29
2.5.7.	Worms attack.....	29
2.5.8.	Probe attack	30
2.5.9.	U2R attack	30
2.5.10.	R2L attack.....	30
2.5.11.	DDoS attack.....	30
2.5.12.	Brute Force	31
2.6.	Instrusion Detection System (IDS)	31
2.7.	Machine Learning	33
2.7.1.	Supervised Learning	33
2.7.2.	Unsupervised Learning.....	33
2.7.3.	Reinforced Learning	33
2.8.	Metode K-Nearest Neighbor (KNN).....	34
2.9.	Dataset.....	36
2.9.1.	Dataset UNSW- NB15.....	36
2.9.2.	Dataset NSL-KDD.....	36
2.9.3.	Dataset ISCX-2012.....	36
2.9.4.	Dataset CIC-IDS-2018	37
2.10.	Google Colaboratory	37

2.11. Confusion Matrix	37
2.11.1. Akurasi.....	38
2.11.2. Presisi.....	38
2.11.3. Recall	39
2.11.4. F1-Score.....	39
BAB III METODELOGI PENELITIAN.....	40
3.1. Kerangka Kerja Penelitian	40
3.2. Tahap Persiapan	42
3.3. Kerangka Kerja Metodologi Penelitian.....	43
3.4. Kebutuhan Perangkat Keras dan Perangkat Lunak.....	44
3.5. Persiapan Dataset	45
3.5.1. Dataset UNSW-NB15.....	45
3.5.2. Dataset NSL-KDD.....	47
3.5.3. Dataset ISCX-2012.....	48
3.5.4. Dataset CIC-IDS-2018	50
3.6. Hasil Ekstraksi Data.....	51
3.7. Feature Selection.....	55
3.8. Metode K-Nearest Neighbor (K-NN)	56
3.9. K-Fold Cross Validation	58
3.10. Validasi Hasil	59
3.11. Pengujian Hyperparameter terhadap Metode K-NN.....	60
3.11.1. Pengujian pada K-Fold Cross Validation	61
3.11.2. Pengujian pada hyperparameter Metode K-NN	64
BAB IV HASIL DAN ANALISIS	69
4.1. Hasil Feature Selection	69
4.1.1. Dataset UNSW-NB15.....	69

4.1.2.	Dataset NSL-KDD.....	73
4.1.3.	Dataset ISCX-2012.....	77
4.1.4.	Dataset CIC-IDS-2018	80
4.2.	Penggunaan SMOTE pada dataset.....	84
4.3.	Pengelompokkan dataset berupa data training dan testing	86
4.4.	Validasi Hasil	87
4.4.1.	Validasi Hasil Dataset UNSW-NB15.....	87
4.4.2.	Validasi Hasil Dataset NSL-KDD.....	125
4.4.3.	Validasi Hasil Dataset ISCX-2012	160
4.4.4.	Validasi Hasil Dataset CIC-IDS-2018.....	195
4.5.	Evaluasi Hasil Rata-Rata Validasi Pada Cross Validation	232
4.6.	Evaluasi Hasil Seluruh Skenario Validasi Pada K-Nearest Neighbor ..	236
4.7.	Analisa Hasil Penelitian	240
4.8.	Perbandingan Terhadap Peneliti Terdahulu	249
BAB V	KESIMPULAN DAN SARAN	250
5.1.	Kesimpulan	250
5.2.	Saran.....	251
DAFTAR PUSTAKA	252

DAFTAR GAMBAR

Gambar 2. 1 Intrusion Detection System	32
Gambar 2. 2 Metode KNN	34
Gambar 3. 1 Kerangka Kerja Penelitian.....	41
Gambar 3. 2 Tahapan Persiapan	42
Gambar 3. 3 Metodologi Penelitian.....	44
Gambar 3. 4 Arsitektur Jaringan pada dataset UNSW-NB15	46
Gambar 3. 5 Tampilan dataset pada UNSW-NB15.....	47
Gambar 3. 6 Tampilan dataset pada NSL-KDD.....	48
Gambar 3. 7 Arsitektur Jaringan pada dataset ISCX-12	48
Gambar 3. 8 Tampilan dataset pada ISCX-2012.....	50
Gambar 3. 9 Topology jaringan dataset pada CIC-IDS-2018	50
Gambar 3. 10 Tampilan dataset pada CIC-IDS-2018.....	51
Gambar 3. 11 Ilustrasi Feature Selection	55
Gambar 3. 12 Flowchart seleksi fitur pada dataset.....	56
Gambar 3. 13 Arsitektur K-Nearest Neighbor (K-NN).....	57
Gambar 3. 14 Arsitektur K-Fold Cross Validation	58
Gambar 3. 15 Kerangka Kerja Deteksi Menggunakan K-NN.....	60
Gambar 4. 1 Grafik korelasi dari dataset UNSW-NB15	70
Gambar 4. 2 Visualisasi Nilai Korelasi UNSW-NB15.....	72
Gambar 4. 3 Grafik korelasi dari dataset NSL-KDD	73
Gambar 4. 4 Visualiasi Nilai Korelasi NSL-KDD	76
Gambar 4. 5 Grafik korelasi dari dataset ISCX-2012	77
Gambar 4. 6 Visualiasi Nilai Korelasi ISCX-2012	79
Gambar 4. 7 Grafik korelasi dari dataset CIC-IDS-2018	80
Gambar 4. 8 Visualiasi Nilai Korelasi CIC-IDS-2018.....	83
Gambar 4. 9 Grafik Smote pada Dataset UNSW-NB15	84
Gambar 4. 10 Grafik Smote pada Dataset NSL-KDD.....	85
Gambar 4. 11 Grafik Smote pada Dataset ISCX-2012.....	85
Gambar 4. 12 Grafik Smote pada Dataset CIC-IDS-2018	86
Gambar 4. 13 Contoh pembagian data training dan data testing.....	87

Gambar 4. 14	Grafik Accuracy Cross Validation UNSW-NB15 rasio 20:80.....	89
Gambar 4. 15	Confusion matrix UNSW-NB15 pada rasio 20:80.....	90
Gambar 4. 16	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 20:80.....	91
Gambar 4. 17	Grafik kurva ROC UNSW-NB15 pada rasio 20:80	92
Gambar 4. 18	Grafik Accuracy Cross Validation UNSW-NB15 rasio 30:70.....	94
Gambar 4. 19	Confusion matrix UNSW-NB15 pada rasio 30:70	95
Gambar 4. 20	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 30:70.....	97
Gambar 4. 21	Grafik kurva ROC UNSW-NB15 pada rasio 30:70	98
Gambar 4. 22	Grafik Accuracy Cross Validation UNSW-NB15 rasio 40:60....	100
Gambar 4. 23	Confusion matrix UNSW-NB15 pada rasio 40:60	101
Gambar 4. 24	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 40:60.....	102
Gambar 4. 25	Grafik kurva ROC UNSW-NB15 pada rasio 40:60	103
Gambar 4. 26	Grafik Accuracy Cross Validation UNSW-NB15 rasio 50:50....	105
Gambar 4. 27	Confusion matrix UNSW-NB15 pada rasio 50:50	106
Gambar 4. 28	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 50:50.....	108
Gambar 4. 29	Grafik kurva ROC UNSW-NB15 pada rasio 50:50	109
Gambar 4. 30	Grafik Accuracy Cross Validation UNSW-NB15 rasio 60:40....	110
Gambar 4. 31	Confusion matrix UNSW-NB15 pada rasio 60:40	111
Gambar 4. 32	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 60:40.....	113
Gambar 4. 33	Grafik kurva ROC UNSW-NB15 pada rasio 60:40	114
Gambar 4. 34	Visualisasi Grafik Accuracy Cross Validation 70:30.....	116
Gambar 4. 35	Confusion matrix UNSW-NB15 pada rasio data 70:30	117
Gambar 4. 36	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 70:30.....	118
Gambar 4. 37	Grafik kurva ROC UNSW-NB15 pada rasio 70:30	119
Gambar 4. 38	Grafik Accuracy Cross Validation UNSW-NB15 80:20.....	121
Gambar 4. 39	Confusion matrix UNSW-NB15 pada rasio 80:20	122
Gambar 4. 40	Grafik Kurva Presisi-Recall UNSW-NB15 pada rasio 80:20.....	123
Gambar 4. 41	Grafik kurva ROC UNSW-NB15 pada rasio 80:20	124
Gambar 4. 42	Grafik Accuracy Cross Validation NSL-KDD rasio 20:80	126
Gambar 4. 43	Confusion matrix NSL-KDD pada rasio 20:80	127
Gambar 4. 44	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 20:80.....	128
Gambar 4. 45	Grafik kurva ROC NSL-KDD pada rasio 20:80.....	129

Gambar 4. 46	Grafik Accuracy Cross Validation NSL-KDD rasio 30:70	131
Gambar 4. 47	Confusion matrix NSL-KDD pada rasio 30:70	132
Gambar 4. 48	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 30:70.....	133
Gambar 4. 49	Grafik kurva ROC NSL-KDD pada rasio 30:70.....	134
Gambar 4. 50	Grafik Accuracy Cross Validation NSL-KDD rasio 40:60	136
Gambar 4. 51	Confusion matrix NSL-KDD pada rasio 40:60	137
Gambar 4. 52	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 40:60.....	138
Gambar 4. 53	Grafik kurva ROC NSL-KDD pada rasio 40:60.....	139
Gambar 4. 54	Grafik Accuracy Cross Validation NSL-KDD rasio 50:50	141
Gambar 4. 55	Confusion matrix NSL-KDD pada rasio 50:50	142
Gambar 4. 56	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 50:50.....	143
Gambar 4. 57	Grafik kurva ROC NSL-KDD pada rasio 50:50.....	144
Gambar 4. 58	Grafik Accuracy Cross Validation NSL-KDD rasio 60:40	146
Gambar 4. 59	Confusion matrix NSL-KDD pada rasio 60:40	147
Gambar 4. 60	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 60:40.....	148
Gambar 4. 61	Grafik kurva ROC NSL-KDD pada rasio 60:40.....	149
Gambar 4. 62	Grafik Accuracy Cross Validation NSL-KDD rasio 70:30	151
Gambar 4. 63	Confusion matrix NSL-KDD pada rasio 70:30	152
Gambar 4. 64	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 70:30.....	153
Gambar 4. 65	Grafik kurva ROC NSL-KDD pada rasio 70:30.....	154
Gambar 4. 66	Grafik Accuracy Cross Validation NSL-KDD rasio 80:20	156
Gambar 4. 67	Confusion matrix NSL-KDD pada rasio 80:20	157
Gambar 4. 68	Grafik Kurva Presisi-Recall NSL-KDD pada rasio 80:20.....	158
Gambar 4. 69	Grafik kurva ROC NSL-KDD pada rasio 80:20.....	159
Gambar 4. 70	Grafik Accuracy Cross Validation ISCX-2012 rasio 20:80	161
Gambar 4. 71	Confusion matrix ISCX-2012 pada rasio 20:80	162
Gambar 4. 72	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 20:80.....	163
Gambar 4. 73	Grafik kurva ROC ISCX-2012 pada rasio 20:80.....	164
Gambar 4. 74	Grafik Accuracy Cross Validation ISCX-2012 rasio 30:70	166
Gambar 4. 75	Confusion matrix ISCX-2012 pada rasio 30:70	167
Gambar 4. 76	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 30:70.....	168
Gambar 4. 77	Grafik kurva ROC ISCX-2012 pada rasio 30:70.....	169

Gambar 4. 78	Grafik Accuracy Cross Validation ISCX-2012 rasio 40:60	171
Gambar 4. 79	Confusion matrix ISCX-2012 pada rasio 40:60	172
Gambar 4. 80	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 40:60.....	173
Gambar 4. 81	Grafik kurva ROC ISCX-2012 pada rasio 40:60.....	174
Gambar 4. 82	Grafik Accuracy Cross Validation ISCX-2012 rasio 50:50	176
Gambar 4. 83	Confusion matrix ISCX-2012 pada rasio 50:50	177
Gambar 4. 84	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 50:50.....	178
Gambar 4. 85	Grafik kurva ROC ISCX-2012 pada rasio 50:50.....	179
Gambar 4. 86	Grafik Accuracy Cross Validation ISCX-2012 rasio 60:40	181
Gambar 4. 87	Confusion matrix ISCX-2012 pada rasio 60:40	182
Gambar 4. 88	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 60:40.....	183
Gambar 4. 89	Grafik kurva ROC ISCX-2012 pada rasio 60:40.....	184
Gambar 4. 90	Grafik Accuracy Cross Validation ISCX-2012 rasio 70:30	186
Gambar 4. 91	Confusion matrix ISCX-2012 pada rasio 70:30	187
Gambar 4. 92	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 70:30.....	188
Gambar 4. 93	Grafik kurva ROC ISCX-2012 pada rasio 70:30.....	189
Gambar 4. 94	Grafik Accuracy Cross Validation ISCX-2012 rasio 80:20	191
Gambar 4. 95	Confusion matrix ISCX-2012 pada rasio 80:20	192
Gambar 4. 96	Grafik Kurva Presisi-Recall ISCX-2012 pada rasio 80:20.....	193
Gambar 4. 97	Grafik kurva ROC ISCX-2012 pada rasio data 80:20	194
Gambar 4. 98	Grafik Accuracy Cross Validation CIC-IDS-2018 pada 20:80 ...	196
Gambar 4. 99	Confusion matrix CIC-IDS-2018 pada rasio 20:80	197
Gambar 4. 100	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 20:80 ..	198
Gambar 4. 101	Grafik kurva ROC CIC-IDS-2018 pada rasio 20:80	199
Gambar 4. 102	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 30:70.	201
Gambar 4. 103	Confusion matrix CIC-IDS-2018 pada rasio 30:70.....	202
Gambar 4. 104	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 30:70 ..	203
Gambar 4. 105	Grafik kurva ROC CIC-IDS-2018 pada rasio 30:70	204
Gambar 4. 106	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 40:60	206
Gambar 4. 107	Confusion matrix CIC-IDS-2018 pada rasio 40:60.....	207
Gambar 4. 108	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 40:60 ..	209
Gambar 4. 109	Grafik kurva ROC CIC-IDS-2018 pada rasio 40:60	210

Gambar 4. 110	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 50:50	212
Gambar 4. 111	Confusion matrix CIC-IDS-2018 pada rasio 50:50	213
Gambar 4. 112	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 50:50 ..	214
Gambar 4. 113	Grafik kurva ROC CIC-IDS-2018 pada rasio 50:50	215
Gambar 4. 114	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 60:40.	217
Gambar 4. 115	Confusion matrix CIC-IDS-2018 pada rasio 60:40	218
Gambar 4. 116	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 60:40 ..	219
Gambar 4. 117	Grafik kurva ROC CIC-IDS-2018 pada rasio 60:40	220
Gambar 4. 118	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 70:30.	222
Gambar 4. 119	Confusion matrix CIC-IDS-2018 pada rasio 70:30	223
Gambar 4. 120	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 70:30 ..	225
Gambar 4. 121	Grafik kurva ROC CIC-IDS-2018 pada rasio 70:30	226
Gambar 4. 122	Grafik Accuracy Cross Validation CIC-IDS-2018 rasio 80:20.	228
Gambar 4. 123	Confusion matrix CIC-IDS-2018 pada rasio 80:20	229
Gambar 4. 124	Grafik Kurva Presisi-Recall CIC-IDS-2018 pada rasio 80:20 ..	230
Gambar 4. 125	Grafik kurva ROC CIC-IDS-2018 pada rasio 80:20	231
Gambar 4. 126	Visualisasi hasil validasi data pada UNSW-NB15	241
Gambar 4. 127	Visualisasi hasil validasi data pada NSL-KDD	243
Gambar 4. 128	Visualisasi hasil validasi data pada ISCX-2012	244
Gambar 4. 129	Visualisasi hasil validasi data pada CIC-IDS-2018	246
Gambar 4. 130	Visualisasi hasil perbandingan terhadap dataset	248

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	8
Tabel 2. 2 Timeline Penelitian Terdahulu	16
Tabel 2. 3 Confusion Matrix	38
Tabel 3. 1 Komponen Perangkat Keras.....	44
Tabel 3. 2 Komponen Perangkat Lunak	45
Tabel 3. 3 Kelompok Fitur UNSW-NB15.....	52
Tabel 3. 4 Kelompok Fitur NSL-KDD.....	52
Tabel 3. 5 Kelompok Fitur ISCX-2012.....	53
Tabel 3. 6 Kelompok Fitur CIC-IDS-2018	54
Tabel 3. 7 Hasil Pengujian K-Fold Cross Validation Dataset UNSW-NB15	61
Tabel 3. 8 Hasil Pengujian K-Fold Cross Validation Dataset NSL-KDD	62
Tabel 3. 9 Hasil Pengujian K-Fold Cross Validation Dataset ISCX-2012.....	62
Tabel 3. 10 Hasil Pengujian K-Fold Cross Validation Dataset CIC-IDS-2018	63
Tabel 3. 11 Pengujian Dataset UNSW-NB15	64
Tabel 3. 12 Pengujian Dataset NSL-KDD	65
Tabel 3. 13 Pengujian Dataset ISCX-2012.....	65
Tabel 3. 14 Pengujian Dataset CIC-IDS-2018	66
Tabel 3. 15 Penggunaan hyperparameter pada metode K-NN	66
Tabel 3. 16 Pembagian Dataset pada UNSW-NB15	67
Tabel 3. 17 Pembagian Dataset pada NSL-KDD	67
Tabel 3. 18 Pembagian Dataset pada ISCX-2012	67
Tabel 3. 19 Pembagian Dataset pada CIC-IDS-2018.....	68
Tabel 4. 1 Nilai Korelasi Fitur pada Dataset UNSW-NB15	71
Tabel 4. 2 Nilai Korelasi Fitur pada Dataset NSL-KDD.....	74
Tabel 4. 3 Nilai Korelasi Fitur pada Dataset ISCX-2012.....	78
Tabel 4. 4 Nilai Korelasi Fitur pada Dataset CIC-IDS-2018	81
Tabel 4. 5 Hasil K-fold Cross validation UNSW-NB15 pada rasio 20:80.....	87
Tabel 4. 6 Hasil Klasifikasi perhitungan UNSW-NB15 rasio 20:80.....	90
Tabel 4. 7 Hasil K-fold Cross validation UNSW-NB15 pada rasio 30:70.....	93
Tabel 4. 8 Hasil Klasifikasi perhitungan UNSW-NB15 rasio 30:70.....	95

Tabel 4. 9	Hasil K-fold Cross validation UNSW-NB15 pada rasio 40:60.....	98
Tabel 4. 10	Hasil Klasifikasi perhitungan UNSW-NB15 rasio 40:60.....	101
Tabel 4. 11	Hasil K-fold Cross validation UNSW-NB15 pada rasio 50:50.....	104
Tabel 4. 12	Hasil Klasifikasi perhitungan UNSW-NB15 rasio 50:50.....	107
Tabel 4. 13	Hasil K-fold Cross validation UNSW-NB15 pada rasio 60:40.....	109
Tabel 4. 14	Hasil Klasifikasi perhitungan UNSW-NB15 rasio 60:40.....	112
Tabel 4. 15	Hasil K-fold Cross validation UNSW-NB15 pada rasio 70:30.....	115
Tabel 4. 16	Hasil Klasifikasi perhitungan UNSW-NB15 rasio 70:30.....	117
Tabel 4. 17	Hasil K-fold Cross validation UNSW-NB15 pada rasio 80:20.....	120
Tabel 4. 18	Hasil Klasifikasi perhitungan UNSW-NB15 rasio 80:20.....	122
Tabel 4. 19	Hasil K-fold Cross validation NSL-KDD pada rasio 20:80.....	125
Tabel 4. 20	Hasil Klasifikasi perhitungan NSL-KDD rasio 20:80.....	127
Tabel 4. 21	Hasil K-fold Cross validation NSL-KDD pada rasio 30:70.....	130
Tabel 4. 22	Hasil Klasifikasi perhitungan NSL-KDD rasio 30:70.....	132
Tabel 4. 23	Hasil K-fold Cross validation NSL-KDD pada rasio 40:60.....	135
Tabel 4. 24	Hasil Klasifikasi perhitungan NSL-KDD rasio 40:60.....	137
Tabel 4. 25	Hasil K-fold Cross validation NSL-KDD pada rasio 50:50.....	140
Tabel 4. 26	Hasil Klasifikasi perhitungan NSL-KDD rasio 50:50.....	142
Tabel 4. 27	Hasil K-fold Cross validation NSL-KDD pada rasio 60:40.....	145
Tabel 4. 28	Hasil Klasifikasi perhitungan NSL-KDD rasio 60:40.....	147
Tabel 4. 29	Hasil K-fold Cross validation NSL-KDD pada rasio 70:30.....	150
Tabel 4. 30	Hasil Klasifikasi perhitungan NSL-KDD rasio 70:30.....	152
Tabel 4. 31	Hasil K-fold Cross validation NSL-KDD pada rasio 80:20.....	155
Tabel 4. 32	Hasil Klasifikasi perhitungan NSL-KDD rasio 80:20.....	157
Tabel 4. 33	Hasil K-fold Cross validation ISCX-2012 pada rasio 20:80.....	160
Tabel 4. 34	Hasil Klasifikasi perhitungan ISCX-2012 rasio 20:80.....	162
Tabel 4. 35	Hasil K-fold Cross validation ISCX-2012 pada rasio 30:70.....	165
Tabel 4. 36	Hasil Klasifikasi perhitungan ISCX-2012 rasio 30:70.....	167
Tabel 4. 37	Hasil K-fold Cross validation ISCX-2012 pada rasio 40:60.....	170
Tabel 4. 38	Hasil Klasifikasi perhitungan ISCX-2012 rasio 40:60.....	172
Tabel 4. 39	Hasil K-fold Cross validation ISCX-2012 pada rasio 50:50.....	175
Tabel 4. 40	Hasil Klasifikasi perhitungan ISCX-2012 rasio 50:50.....	177

Tabel 4. 41	Hasil K-fold Cross validation ISCX-2012 pada rasio 60:40.....	180
Tabel 4. 42	Hasil Klasifikasi perhitungan ISCX-2012 rasio 60:40.....	182
Tabel 4. 43	Hasil K-fold Cross validation ISCX-2012 pada rasio 70:30.....	185
Tabel 4. 44	Hasil Klasifikasi perhitungan ISCX-2012 rasio 70:30.....	187
Tabel 4. 45	Hasil K-fold Cross validation ISCX-2012 pada rasio 80:20.....	190
Tabel 4. 46	Hasil Klasifikasi perhitungan ISCX-2012 rasio 80:20.....	192
Tabel 4. 47	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 20:80	195
Tabel 4. 48	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 20:80	197
Tabel 4. 49	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 30:70	200
Tabel 4. 50	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 30:70	202
Tabel 4. 51	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 40:60	205
Tabel 4. 52	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 40:60	208
Tabel 4. 53	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 50:50	211
Tabel 4. 54	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 50:50	213
Tabel 4. 55	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 60:40	216
Tabel 4. 56	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 60:40	218
Tabel 4. 57	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 70:30	221
Tabel 4. 58	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 70:30	224
Tabel 4. 59	Hasil K-fold Cross validation CIC-IDS-2018 pada rasio 80:20	227
Tabel 4. 60	Hasil Klasifikasi perhitungan CIC-IDS-2018 rasio 80:20	229
Tabel 4. 61	Akurasi Berdasarkan Cross Validation pada UNSW-NB15	232
Tabel 4. 62	Akurasi Berdasarkan Cross Validation pada NSL-KDD	233
Tabel 4. 63	Akurasi berdasarkan rata-rata cross validation ISCX-2012.....	234
Tabel 4. 64	Akurasi Berdasarkan Cross Validation pada CIC-IDS-2018.....	235
Tabel 4. 65	Hasil pengujian pada dataset UNSW-NB15	236
Tabel 4. 66	Hasil pengujian pada dataset NSL-KDD	237
Tabel 4. 67	Hasil pengujian dataset ISCX-2012	238
Tabel 4. 68	Hasil pengujian pada dataset CIC-IDS-2018	239
Tabel 4. 69	Hasil performa terbaik pengujian pada UNSW-NB15.....	240
Tabel 4. 70	Hasil performa terbaik pengujian pada NSL-KDD.....	242
Tabel 4. 71	Hasil performa terbaik pengujian pada ISCX-2012.....	243
Tabel 4. 72	Hasil performa terbaik pengujian pada CIC-IDS-2018.....	245

Tabel 4. 73 Hasil Perbandingan Terhadap Dataset	247
Tabel 4. 74 Hasil perbandingan terhadap penelitian terkait	249

DAFTAR LAMPIRAN

Lampiran 1. Form Perbaikan.....	257
Lampiran 2. Hasil Cek Plagiat.....	259

BAB I

PENDAHULUAN

1.1. Latar Belakang

Ancaman serangan siber semakin serius bagi organisasi dan individu diseluruh dunia. Serangan siber merupakan jenis serangan yang ditujukan untuk merusak, mengubah, atau mencuri data penting pada komputer atau jaringan komputer. Serangan siber dapat berupa tindakan peretasan, phishing, malware, dan berbagai jenis serangan lainnya [1]. Dampak dari serangan siber dapat menyebabkan kerugian finansial dan reputasi yang signifikan bagi individual atau organisasi yang menjadi korban. Oleh karena itu, penting bagi kita untuk mengembangkan sistem deteksi serangan siber yang efektif demi menjaga keamanan data dan sistem informasi.

Sistem deteksi serangan siber merupakan salah satu aspek penting dalam keamanan siber untuk melindungi sistem informasi dari ancaman serangan siber yang semakin meningkat [2]. Terkait laporan Badan Siber dan Sandi Negara (BSSN) mempublikasikan laporan tahunan berjudul “Monitoring Keamanan Siber” untuk tahun 2022. Laporan tersebut mengungkapkan bahwa terdapat lebih dari 900 juta atau tepatnya 973.952.290 anomali traffic atau serangan siber (*cyber attack*) yang terjadi diseluruh wilayah Indonesia sepanjang tahun 2022 [3]. Permintaan akan keamanan siber dan perlindungan terhadap anomali siber dan berbagai macam serangan, seperti akses tidak sah, Denial-Of-Service (DOS), phishing, malware, botnet, spyware, worm dan sebagainya meningkat secara dramatis dalam beberapa hari terakhir [4]. Oleh karena itu dibutuhkan aplikasi yang mampu menganalisis data cerdas untuk memberikan pola perlindungan terhadap serangan dimasa depan.

Semakin meningkat serangan siber, membuat banyak para penelitian mencari cara bagaimana untuk mengatasi masalah serangan tersebut. Salah satunya adalah Intrusion Detection System (IDS). IDS merupakan suatu alat yang digunakan untuk mendeteksi berbagai ancaman keamanan di komputer jaringan [5]. IDS akan memantau sebuah peristiwa yang akan terjadi didalam suatu sistem dan

mengambil tindakan pada peristiwa apakah disebut serangan atau tindakan yang sah dalam suatu sistem.

Untuk meningkatkan kinerja IDS, para peneliti telah mengeksplorasi penggunaan Artificial Intelligence (AI) dan lebih khusus lagi penerapan teknik berbasis Machine Learning (ML). ML merupakan cabang AI yang memberdayakan berbagai sistem kecerdasan yang dapat membuat komputer memiliki kemampuan dan kapasitas untuk memperbaiki proses pengambilan keputusan tanpa harus diprogram secara eksplisit [6]. Pada penelitian ini metode ML yang akan digunakan adalah ML berbasis Supervised Learning yakni K-Nearest Neighbor (KNN).

Metode K-NN merupakan metode klasifikasi yang mengukur jarak antara data yang diberikan dan data yang telah diklasifikasikan sebelumnya. Algoritma K-NN merupakan teknik nonparametrik yang digunakan untuk melakukan klasifikasi dan regresi. Dalam regresi K-NN, algoritma KNN digunakan untuk memperkirakan variabel label kontinu [7]. Algoritma KNN merupakan suatu algoritma yang sering digunakan dalam klasifikasi, tujuan dari algoritma ini yaitu untuk mengklasifikasi suatu objek baru. Objek baru akan diklasifikasikan dengan cara mencari jarak terdekat ke objek pelatihan yang sudah ada berdasarkan pada rumus Euclidean distance. Setelah itu, objek baru akan diklasifikasikan ke kelas yang paling banyak ditemukan diantara k tetangga terdekatnya [8] .

Serangan siber dapat dibagi menjadi berbagai kategori, dan untuk setiap kategori serangan, perlu digunakan teknik khusus untuk mendeteksinya. Oleh karena itu, didalam penelitian ini akan dikembangkan sistem deteksi *multi-classification* serangan siber dengan menggunakan metode K-NN. sistem deteksi ini akan mampu mendeteksi berbagai jenis serangan siber dengan menggunakan teknik yang berbeda untuk setiap jenis serangan [9]. Penelitian ini bertujuan untuk mengembangkan sistem deteksi serangan siber yang efektif dan dapat diandalkan. Diharapkan hasil penelitian ini dapat membantu organisasi dan individu dalam melindungi sistem informasi mereka dari ancaman serangan siber yang semakin meningkat.

Pada Penelitian [10] berdasarkan hasil yang didapatkan dari pemilihan fitur menggunakan algoritma genetika memiliki nilai fitness sebesar 84,171% , dan fitur yang terpilih adalah 18 dari 41 fitur yang ada. Hasil evaluasi pengujian dengan

menggunakan dataset KDD-CUP 99 dengan metode K-NN, diperoleh akurasi data training 99,89% dan data testing sebesar 97,54%. Akurasi tersebut didapat pada parameter $k = 5$ sedangkan $k = 7$ rata-rata akurasi yang diperoleh adalah sebesar 97,52%. Akurasi hasil pengujian manual pada evaluasi $k = 1, 3, 5, 7$, dan 9 dari 1000, 5000, 10000, 15000, dan 30000 didapat akurasi rata-rata 78,57%, 76,40%, 76,86%, 76,71%, dan 77,57%. Parameter tersebut memiliki tingkat akurasi tertinggi sebesar 78,57% pada parameter $k = 1$ dan tingkat akurasi terendah dengan akurasi sebesar 76,40% pada parameter $k = 3$.

Menurut penelitian [11] disebutkan bahwa beberapa algoritma seperti K-NN, Decision Tree, Logistic Regression, dan Random Forest telah. Hasil evaluasi menunjukkan bahwa algoritma K-NN memiliki performa terbaik dibandingkan dengan algoritma lainnya, dengan tingkat akurasi 95,51%, recall 89,42%, precision 89,42%. Sementara untuk hasil 10-fold cross validation, algoritma K-NN memiliki tingkat akurasi sebesar 93,61%, recall 85,05%, precision 85,25 dalam mendeteksi website malicious dan benign.

Berdasarkan pada uraian-uraian tersebut, penulis melakukan penelitian sistem deteksi serangan siber dengan menggunakan metode K-Nearest Neighbor pada dataset UNSW-NB15, NSL-KDD, ISCX-2012 dan CIC-IDS-2018 dan teknik multi-classification diharapkan mampu meningkatkan akurasi dan efektivitas deteksi serangan siber, serta dapat membantu melindungi sistem informasi dari ancaman serangan siber yang semakin meningkat. Itulah yang melatar belakangi penulis mengambil judul "**Optimalisasi Multi-Classification Serangan Cyber Menggunakan Metode K-Nearest Neighbor**". Dikarenakan metode K-NN dianggap memperoleh tingkat akurasi yang baik dalam mengklasifikasikan sebuah data.

1.2. Tujuan

Tujuan dari penulisan Tugas Akhir ini adalah sebagai berikut :

1. Melakukan penerapan metode K-Nearest Neighbor dalam mendeteksi multi-classification serangan siber

2. Mencari tingkat akurasi terbaik, percision, recall dan F1- score dalam mendeteksi serangan siber dengan menggunakan metode K-NN pada sistem *deteksi multi-classification* serangan siber.

1.3. Manfaat

Adapun manfaat dari peulisan Tugas Akhir ini adalah sebagai berikut:

1. Dapat menerapkan metode K-Nearest Neighbor dalam mendeteksi multi-classification serangan siber.
2. Dapat memberikan informasi mengenai akurasi,precision, recall, dan F1-score menggunakan metode K-Nearest Neighbor dalam melakukan deteksi *multi-classification* serangan siber.

1.4. Perumusan Masalah

Adapun beberapa point rumusan masalah yang didapatkan dalam penulisan Tugas Akhir ini adalah sebagai berikut:

1. Bagaimana cara melakukan deteksi multi-classification serangan siber dengan menggunakan metode K-Nearest Neighbor?
2. Bagaimana cara mencari tingkat akurasi terbaik, precision, recall dan F1-score dalam mendeteksi *multi-classification* serangan siber dengan menggunakan metode K-Nearest Neighbor?

1.5. Batasan Masalah

Batasan masalah merupakan batasan yang digunakan dalam penelitian agar penelitian tidak melenceng terlalu jauh. Dengan demikian batasan masalah dalam penyusunan Tugas Akhir ini sebagai berikut :

1. Penelitian ini hanya menggunakan dataset UNSW NB-15, NSL-KDD, ISCX-2012 dan CIC.IDS.2018.
2. Penelitian ini menggunakan program bahasa Python.
3. Algoritma klasifikasi yang digunakan dalam penelitian ini adalah Algoritma K-Nearest Neighbor.

4. Penelitian ini hanya menghasilkan output berupa akurasi terbaik, precision, recall, dan F1-score dalam mendeteksi *multi-classification* serangan siber dengan melihat kecocokan author dan tabel.

1.6. Metodologi Penelitian

Metodologi yang diterapkan dalam penyusunan Tugas Akhir ini melalui beberapa tahapan sebagai berikut :

1. Tahap pertama (Studi Pustaka)

Tahap pertama yang akan dilakukan dalam penelitian ini adalah mencari sumber informasi, memahami, serta mempelajari kajian literatur dan referensi seperti buku, artikel yang terkait, maupun jurnal ilmiah yang berkaitan dengan serangan siber dan konsep dari metode K-Nearest Neighbor. Sehingga dapat memberikan penunjang pada metodologi yang akan diterapkan dalam penelitian ini.

2. Tahap Kedua (Perancangan Sistem)

Pada tahap kedua membahas tentang perancangan sistem yang akan dilakukan dalam penelitian ini. Dimana, tahap ini akan membuat suatu perancangan menggunakan pemodelan simulasi dengan program bahasa Python.

3. Tahap Ketiga (Pengujian)

Pada tahap ketiga ini adalah melakukan pengujian data. Tahap ini dilakukan jika semua sistem telah dibuat dan dikonfigurasi, maka akan dilakukan pengujian sesuai batasan masalah pada penelitian agar mendapatkan hasil yang optimal.

4. Tahap Keempat (Analisa)

Selanjutnya pada tahap ini akan dilakukan analisa dari hasil tahap pengujian. Dimana pada tahap ini akan dilakukan analisa dari proses penelitian yang tujuannya untuk mengetahui kekurangan pada hasil perancangan dan faktor apa yang menjadi penyebabnya. Sehingga, bisa dilakukan pengembangan lagi untuk peneliti selanjutnya.

5. Tahap Kelima (Kesimpulan dan Saran)

Tahap ini merupakan tahapan terakhir, dimana penulis melakukan penarikan kesimpulan yang berdasarkan studi pustaka, hasil perancangan sistem, serta hasil analisa penelitian. Dan memberikan saran untuk peneliti selanjutnya sebagai bahan referensi.

1.7. Sistematika Penulisan

Adapun penyusunan penulisan Tugas Akhir ini disusun menjadi beberapa sub bab yang akan dijelaskan secara rinci mengenai apa saja yang dilakukan oleh penulis pada saat melakukan penelitian. Sistematika penulisan untuk Tugas Akhir ini disusun sebagai berikut:

BAB I PENDAHULUAN

Bab pertama berisi mengenai penjelasan secara sistematis berupa topik penelitian yang meliputi latar belakang, tujuan, manfaat, perumusan masalah, batasan masalah, metodologi yang digunakan dan terakhir mengenai sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab kedua akan menjelaskan dasar teori penelitian mengenai serangan siber, sistem deteksi multi-classification dan metode dari Algoritma K-Nearest Neighbor yang berkaitan langsung dengan penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ketiga berisi uraian secara sistematis bagaimana proses yang dilakukan dalam penelitian ini. Pada bab ini akan dilakukan pengkajian tahapan perancangan sistem dan bagaimana penerapan dari metode penelitian ini.

BAB IV PENGUJIAN DAN ANALISIS

Bab keempat akan menjelaskan hasil dari proses pengujian yang telah dilakukan, serta akan dilakukan analisis dari data yang didapat dari hasil pengujian.

BAB V KESIMPULAN DAN SARAN

Bab terakhir berisi mengenai kesimpulan dan saran dari hasil analisa yang diperoleh berdasarkan penelitian yang telah dilakukan.

DAFTAR PUSTAKA

- [1] M. A. Hasnat and M. Rahnamay-Naeini, "Detecting and locating cyber and physical stresses in smart grids using the k-nearest neighbour analysis of instantaneous correlation of states," *IET Smart Grid*, vol. 4, no. 3, pp. 307–320, 2021, doi: 10.1049/stg2.12030.
- [2] F. Zhao and Q. Tang, "A KNN learning algorithm for collusion-resistant spectrum auction in small cell networks," *IEEE Access*, vol. 6, pp. 45796–45803, 2018, doi: 10.1109/ACCESS.2018.2861840.
- [3] B. Publik, "D E S E M B E R 2 0 2 2 L a P O R a N," no. 70, 2022, [Online]. Available: www.idsirtii.or.id
- [4] I. H. Sarker, "Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects," *Ann. Data Sci.*, 2022, doi: 10.1007/s40745-022-00444-2.
- [5] H. M. Anwer, M. Farouk, and A. Abdel-Hamid, "A framework for efficient network anomaly intrusion detection with features selection," *2018 9th Int. Conf. Inf. Commun. Syst. ICICS 2018*, vol. 2018-Janua, pp. 157–162, 2018, doi: 10.1109/IACS.2018.8355459.
- [6] S. M. Kasongo, "An advanced intrusion detection system for IIoT Based on GA and tree based algorithms," *IEEE Access*, vol. 9, pp. 113199–113212, 2021, doi: 10.1109/ACCESS.2021.3104113.
- [7] K. Veena, K. Meena, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "C SVM Classification and KNN Techniques for Cyber Crime Detection," *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/3640017.
- [8] Okfalisa, I. Gazalba, Mustakim, and N. G. I. Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-Janua, pp. 294–298, 2018, doi: 10.1109/ICITISEE.2017.8285514.
- [9] B. B. Rao and K. Swathi, "Fast kNN Classifiers for Network Intrusion

- Detection System,” *Indian J. Sci. Technol.*, vol. 10, no. 14, pp. 1–10, 2017, doi: 10.17485/ijst/2017/v10i14/93690.
- [10] M. A. Fauzi, A. T. Hanuranto, and C. Setianingsih, “Intrusion Detection System using Genetic Algorithm and K-NN Algorithm on Dos Attack,” *2020 2nd Int. Conf. Cybern. Intell. Syst. ICORIS 2020*, pp. 3–8, 2020, doi: 10.1109/ICORIS50180.2020.9320822.
- [11] G. A. Sandag, J. Leopold, and V. F. Ong, “Klasifikasi Malicious Websites Menggunakan Algoritma K-NN Berdasarkan Application Layers dan Network Characteristics,” *CogITO Smart J.*, vol. 4, no. 1, pp. 37–45, 2018, doi: 10.31154/cogito.v4i1.100.37-45.
- [12] K. Atefi, H. Hashim, and M. Kassim, “Anomaly analysis for the classification purpose of intrusion detection system with K-nearest neighbors and deep neural network,” *Proceeding - 2019 IEEE 7th Conf. Syst. Process Control. ICSPC 2019*, no. December, pp. 269–274, 2019, doi: 10.1109/ICSPC47137.2019.9068081.
- [13] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, “CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System,” *IEEE Access*, vol. 10, no. August, pp. 99837–99849, 2022, doi: 10.1109/ACCESS.2022.3206425.
- [14] M. Shorfuzzaman, “Detection of cyber attacks in IoT using tree-based ensemble and feedforward neural network,” *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2020-Octob, pp. 2601–2606, 2020, doi: 10.1109/SMC42975.2020.9283443.
- [15] L. Zhiqiang, G. Mohi-Ud-Din, L. Bing, L. Jianchao, Z. Ye, and L. Zhijun, “Modeling Network Intrusion Detection System Using Feed-Forward Neural Network Using UNSW-NB15 Dataset,” *Proc. 2019 7th Int. Conf. Smart Energy Grid Eng. SEGE 2019*, pp. 299–303, 2019, doi: 10.1109/SEGE.2019.8859773.
- [16] M. Zeeshan *et al.*, “Protocol-Based Deep Intrusion Detection for DoS and DDoS Attacks Using UNSW-NB15 and Bot-IoT Data-Sets,” *IEEE Access*, vol. 10, pp. 2269–2283, 2022, doi: 10.1109/ACCESS.2021.3137201.
- [17] D. Jing and H. B. Chen, “SVM based network intrusion detection for the

- UNSW-NB15 dataset,” *Proc. Int. Conf. ASIC*, pp. 1–4, 2019, doi: 10.1109/ASICON47005.2019.8983598.
- [18] Z. R. Tembusai, H. Mawengkang, and M. Zarlis, “K-Nearest Neighbor with K-Fold Cross Validation and Analytic Hierarchy Process on Data Classification,” *Int. J. Adv. Data Inf. Syst.*, vol. 2, no. 1, pp. 1–8, 2021, doi: 10.25008/ijadis.v2i1.1204.
- [19] S. M. Kasongo and Y. Sun, “A deep learning method with filter based feature engineering for wireless intrusion detection system,” *IEEE Access*, vol. 7, pp. 38597–38607, 2019, doi: 10.1109/ACCESS.2019.2905633.
- [20] S. N. Nguyen, V. Q. Nguyen, J. Choi, and K. Kim, “Design and implementation of intrusion detection system using convolutional neural network for DoS detection,” *ACM Int. Conf. Proceeding Ser.*, pp. 34–38, 2018, doi: 10.1145/3184066.3184089.
- [21] 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, 2019, doi: 10.1016/j.knosys.2018.08.036.
- [22] M. D. Hossain, H. Ochiai, F. Doudou, and Y. Kadobayashi, “SSH and FTP brute-force attacks detection in computer networks: Lstm and machine learning approaches,” *2020 5th Int. Conf. Comput. Commun. Syst. ICCCS 2020*, pp. 491–497, 2020, doi: 10.1109/ICCCS49078.2020.9118459.
- [23] A. A. E. B. Donkol, A. G. Hafez, A. I. Hussein, and M. M. Mabrook, “Optimization of Intrusion Detection Using Likely Point PSO and Enhanced LSTM-RNN Hybrid Technique in Communication Networks,” *IEEE Access*, vol. 11, no. February, pp. 9469–9482, 2023, doi: 10.1109/ACCESS.2023.3240109.
- [24] N. Gandhi, J. Patel, R. Sisodiya, N. Doshi, and S. Mishra, “A CNN-BiLSTM based Approach for Detection of SQL Injection Attacks,” *Proc. 2nd IEEE Int. Conf. Comput. Intell. Knowl. Econ. ICCIKE 2021*, pp. 378–383, 2021, doi: 10.1109/ICCIKE51210.2021.9410675.
- [25] A. Shah, S. Clachar, M. Minimair, and D. Cook, “Building multiclass classification baselines for anomaly-based network intrusion detection systems,” *Proc. - 2020 IEEE 7th Int. Conf. Data Sci. Adv. Anal. DSAA 2020*,

- pp. 759–760, 2020, doi: 10.1109/DSAA49011.2020.00102.
- [26] A. R. Syarif and W. Gata, “Intrusion detection system using hybrid binary PSO and K-nearest neighborhood algorithm,” *Proc. 11th Int. Conf. Inf. Commun. Technol. Syst. ICTS 2017*, vol. 2018-Janua, pp. 181–186, 2018, doi: 10.1109/ICTS.2017.8265667.
- [27] B. Xu, S. Chen, H. Zhang, and T. Wu, “Incremental k-NN SVM method in intrusion detection,” *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 2017-Novem, pp. 712–717, 2018, doi: 10.1109/ICSESS.2017.8343013.
- [28] G. Karatas, O. Demir, and O. K. Sahingoz, “Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset,” *IEEE Access*, vol. 8, pp. 32150–32162, 2020, doi: 10.1109/ACCESS.2020.2973219.
- [29] A. Luo, W. Huang, and W. Fan, “A CNN-based Approach to the Detection of SQL Injection Attacks,” *Proc. - 18th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2019*, pp. 320–324, 2019, doi: 10.1109/ICIS46139.2019.8940196.
- [30] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, “A detailed investigation and analysis of using machine learning techniques for intrusion detection,” *IEEE Commun. Surv. Tutorials*, vol. 21, no. 1, pp. 686–728, 2019, doi: 10.1109/COMST.2018.2847722.
- [31] J. M. Biju, N. Gopal, and A. J. Prakash, “Cyber Attacks and Its Different Types,” *Int. Res. J. Eng. Technol.*, vol. 6, no. 3, pp. 4849–4852, 2019, [Online]. Available: <https://www.irjet.net/archives/V6/i3/IRJET-V6I31244.pdf>
- [32] Y. Y. Aung and M. Myat Min, “Hybrid Intrusion Detection System using K-means and K-Nearest Neighbors Algorithms,” *Proc. - 17th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2018*, pp. 34–38, 2018, doi: 10.1109/ICIS.2018.8466537.
- [33] Y. Xin *et al.*, “Machine Learning and Deep Learning Methods for Cybersecurity,” *IEEE Access*, vol. 6, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.
- [34] Y. Liao and V. R. Vemuri, “Classifier for Intrusion,” *Comput. Secur.*, vol. 21, no. 5, pp. 439–448, 2002.

- [35] V. Hnamte, H. Nhung-Nguyen, J. Hussain, and Y. Hwa-Kim, “A Novel Two-Stage Deep Learning Model for Network Intrusion Detection: LSTM-AE,” *IEEE Access*, vol. 11, pp. 37131–37148, 2023, doi: 10.1109/ACCESS.2023.3266979.
- [36] A. Singh, M. Singh, and K. Kumar, “A Hybrid Method for Intrusion Detection Using SVM and k-NN,” *Lect. Notes Networks Syst.*, vol. 175, no. February, pp. 119–126, 2021, doi: 10.1007/978-3-030-67187-7_13.
- [37] A. R. Sonule, M. Kalla, A. Jain, and D. S. Chouhan, “Unsw-Nb15 Dataset and Machine Learning Based Intrusion Detection Systems,” *Int. J. Eng. Adv. Technol.*, vol. 9, no. 3, pp. 2638–2648, 2020, doi: 10.35940/ijeat.c5809.029320.
- [38] M. Canesche, L. Bragança, O. P. V. Neto, J. A. Nacif, and R. Ferreira, “Google Colab CAD4U: Hands-on cloud laboratories for digital design,” *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2021-May, no. Vcd, 2021, doi: 10.1109/ISCAS51556.2021.9401151.
- [39] I. Markoulidakis, G. Kopsiaftis, I. Rallis, and I. Georgoulas, “Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem,” *ACM Int. Conf. Proceeding Ser.*, no. Cx, pp. 412–419, 2021, doi: 10.1145/3453892.3461323.
- [40] T. Kim and W. Pak, “Hybrid Classification for High-Speed and High-Accuracy Network Intrusion Detection System,” *IEEE Access*, vol. 9, pp. 83806–83817, 2021, doi: 10.1109/ACCESS.2021.3087201.