

**IMPLEMENTASI ALGORITMA KLASIFIKASI (RANDOM
FOREST, XGBOOST, LIGHTGBM, CATBOOST) DALAM
MENDETEKSI SERANGAN JARINGAN INTERNET**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh
PURNAMASARI
21.11. 3850

Kepada
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

**IMPLEMENTASI ALGORITMA KLASIFIKASI (RANDOM
FOREST, XGBOOST, LIGHTGBM, CATBOOST) DALAM
MENDETEKSI SERANGAN JARINGAN INTERNET**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh
PURNAMASARI
21.11.3850

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2025

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA KLASIFIKASI (RANDOM FOREST,
XGBOOST, LIGHTGBM, DAN CATBOOST) DALAM MENDETEKSI
SERANGAN JARINGAN INTERNET**

yang disusun dan diajukan oleh

Purnamasari

21.11.3850

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 juli 2025

Dosen Pembimbing,



Subektiningsih, S.Kom., M.Kom.
NIK. 190362413

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA KLASIFIKASI (RANDOM FOREST, XGBOOST,
LIGHTGBM, DAN CATBOOST) DALAM MENDETEKSI SERANGAN JARINGAN
INTERNET**

yang disusun dan diajukan oleh

PURNAMASARI

21.11.3850

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Rumini, S.Kom., M.Kom.
NIK. 190302246



Wiwi Widayani, S.Kom., M.Kom.
NIK. 190302272



Subektiingsih, S.Kom., M.Kom.
NIK. 190302413



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Purnamasari
NIM : 21.11.3850

Menyatakan bahwa Skripsi dengan judul berikut:

Implementasi Algoritma Klasifikasi (Random Forest, XGBoost, LightGBM, Dan CatBoost) Dalam Mendeteksi Serangan Jaringan Internet

Dosen Pembimbing: Subektiningsih, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 Juli 2025

Yang Menyatakan,



The image shows an official stamp of Universitas AMIKOM Yogyakarta. The stamp is rectangular and contains the university's logo, the name 'UNIVERSITAS AMIKOM YOGYAKARTA', and the text 'MUTUAL TEMPER'. Overlaid on the stamp is a handwritten signature in black ink.

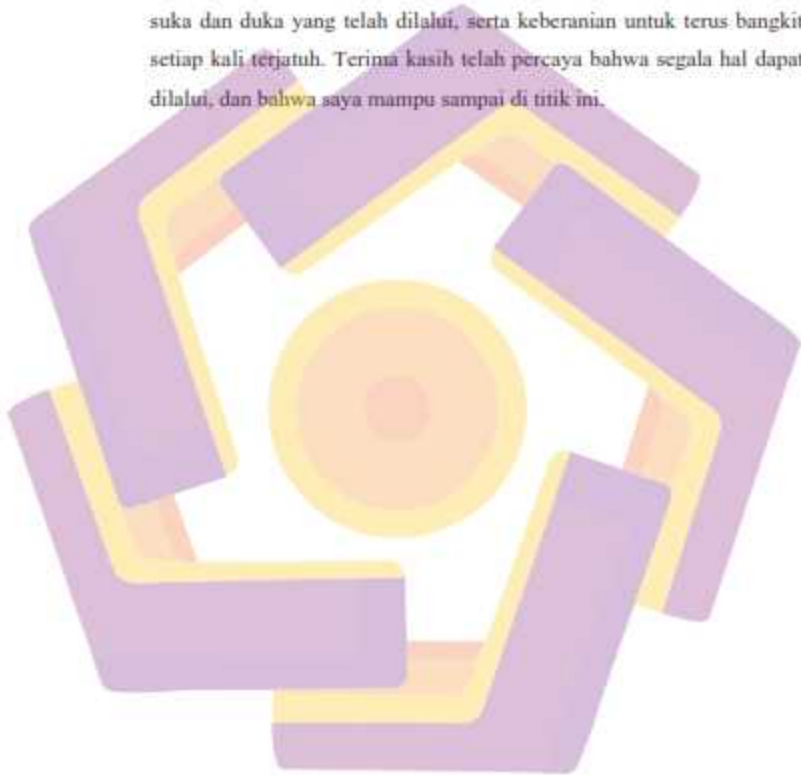
Purnamasari

HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kepada Allah SWT atas segala limpahan rahmat dan ridhoNya yang telah memberikan kesehatan, kelancaran dan kekuatan. Atas segala karunia serta kemudahan yang Engkau berikan akhirnya skripsi ini dapat terselesaikan. Selain itu penulis juga berterimakasih kepada orang yang sangat berarti dalam pembuatan skripsi ini:

1. Kepada kedua orang tua saya, Ayah dan Mamak yang telah mendukung penuh perkuliahan saya selama ini. Doa, dukungan dan kerja keras kalian yang sangat berarti. Juga kepada dua adik laki-laki saya Helmi dan Arga yang menjadi semangat saya untuk terus melangkah.
2. Dosen Pembimbing saya ibu Subektiningsih yang telah membimbing dan memperlancar proses saya dalam skripsi ini.
3. Seluruh Dosen di Universitas Amikom Yogyakarta yang telah mengajar matakuliah yang saya ambil selama ini, yang telah berjasa mengajari dan memberi ilmu sehingga saya dapat berkembang lagi.
4. Almarhum teman saya Restu Fauzi yang telah menemani dan membantu dalam proses perkuliahan saya.
5. Sahabat saya Arstitan yang selalu ada dan mendukung saya selama ini dari SMP sampai sekarang.
6. Sahabat saya Intan dan Husni yang menemani proses saat saya sedang terpuruk hingga dapat melanjutkan kembali skripsi ini.
7. Sahabat saya Uli yang selalu mendukung, memberi motivasi dan banyak membantu saya dari SMA sampai sekarang.
8. Sahabat saya Bella yang selalu senantiasa mendukung proses dan membantu saya dalam pembuatan skripsi ini.
9. Teman dekat saya Erman dan Cahyo yang menemani proses pembuatan skripsi ini sehingga bisa berjalan lancar dan selalu memberikan saya dukungan penuh.
10. Fatan, Rais, Mafi dan teman-teman kampus saya yang sudah banyak membantu saya dalam proses skripsi ini.

11. Orang-orang yang datang menaruh luka pada masanya dan meninggalkan banyak hal yang membuat saya belajar dan membuat saya menjadi lebih kuat.
12. Dan terakhir, saya persembahkan skripsi ini untuk diri saya sendiri. Terimakasih, untuk selalu kuat dan tidak pernah menyerah. Untuk segala suka dan duka yang telah dilalui, serta keberanian untuk terus bangkit setiap kali terjatuh. Terima kasih telah percaya bahwa segala hal dapat dilalui, dan bahwa saya mampu sampai di titik ini.



KATA PENGANTAR

Puji Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, karunia, dan ridho-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul “Perbandingan Kinerja Algoritma Klasifikasi Dalam Mendeteksi Serangan Jaringan Internet”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Dalam proses penyusunan skripsi ini, penulis telah menerima banyak bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta, yang telah memberikan dukungan, fasilitas, dan kesempatan dalam menyelesaikan studi serta penyusunan skripsi ini di lingkungan akademik yang kondusif dan inspiratif.
2. Ibu Subektiningsih, M.Kom., selaku dosen pembimbing, yang telah dengan membimbing dan memberikan arahan selama proses penyusunan skripsi ini.
3. Tim dosen penguji skripsi yang telah memberikan masukan dan evaluasi yang sangat berharga demi kesempurnaan karya ilmiah ini.
4. Kedua orang tua penulis, Habiburrahman dan Siti Aminah, atas doa, dukungan moral, serta kasih sayang yang tiada henti selama penulis menjalani pendidikan.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis terbuka terhadap segala bentuk kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca serta pihak-pihak yang membutuhkan.

Yogyakarta, 21 Juli 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xiii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penelitian.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	15
2.2.1 Jaringan Komputer.....	15

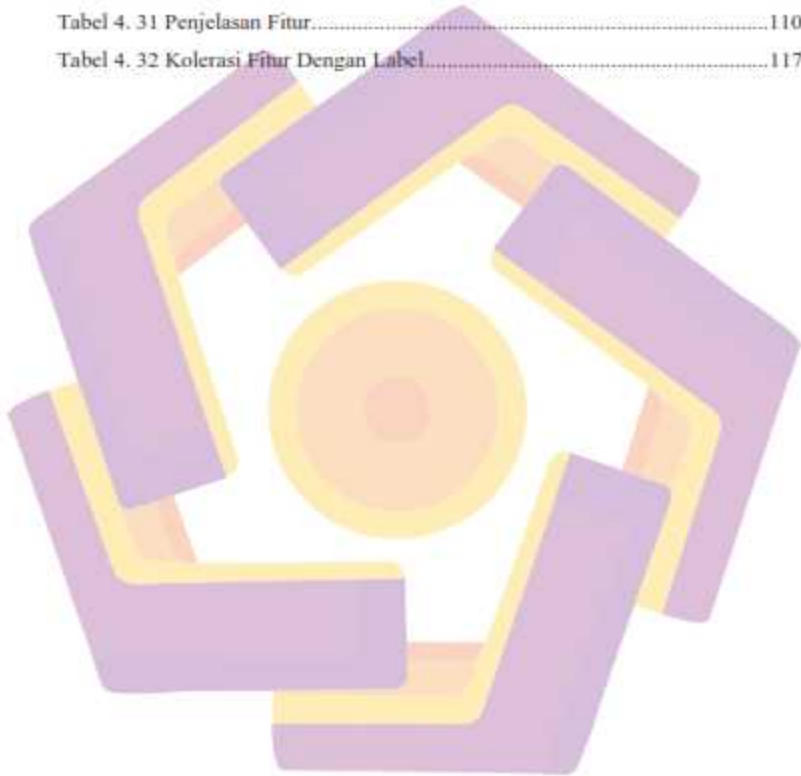
2.2.2	Kecamanan Jaringan	15
2.2.3	Serangan Jaringan	15
2.2.4	Machine Learning	19
2.2.5	SMOTETomek	24
2.2.6	Random Forest	25
2.2.7	XGBoost (<i>Extreme Gradient Boosting</i>)	26
2.2.8	LightGBM	26
2.2.9	CatBoost	26
2.2.10	Python	27
2.2.11	Overfitting & Underfitting	27
BAB III METODE PENELITIAN		29
3.1	Objek Penelitian	29
3.2	Alur Penelitian	29
3.2.1	Pengumpulan Data	30
3.2.2	Exploration Data Analysis (EDA)	30
3.2.3	Pre-pemrosesan Data	32
3.2.4	Modelling	34
3.2.7	Perbandingan Algoritma	43
3.2.8	Deployment	43
3.3	Alat dan Bahan	43
3.3.1	Data Penelitian	43
3.3.2	Alat	44
BAB IV HASIL DAN PEMBAHASAN		45
4.1	Pengambilan Data	45
4.2	<i>Exploration Data Analysis (EDA)</i>	46

4.3	Preprocessing	51
4.3.1	Encoding	52
4.3.2	SMOTETomek	56
4.4	Modelling	62
4.4.1	Random Forest	62
4.4.2	XGBoost (EXtreme Gradient Boosting)	66
4.4.3	LightGBM	68
4.4.4	CatBoost	71
4.2	Evaluasi	74
4.5.1	Random Forest	75
4.5.2	XGBoost	83
4.5.3	LightGBM	90
4.5.4	CatBoost	95
4.6	Perbandingan Hasil Algoritma	102
4.7	Deployment	107
BAB V PENUTUP		121
5.1	Kesimpulan	121
5.2	Suran	121
REFERENSI		123
LAMPIRAN		127

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	10
Tabel 3. 1 Distribution Label	30
Tabel 3. 2 Confusion Matrix	36
Tabel 3. 3 Hardware	44
Tabel 4. 1 Code Pengambilan Data	45
Tabel 4. 2 Code EDA	47
Tabel 4. 3 Code Tahap Encoding	52
Tabel 4. 4 Code Tahap SMOTETomek	57
Tabel 4. 5 Code Tahap Visualisasi SMOTETomek	59
Tabel 4. 6 Code Model Random Forest	62
Tabel 4. 7 Code Model XGBoost	66
Tabel 4. 8 Code Model LightGBM	68
Tabel 4. 9 Code Model CatBoost	71
Tabel 4. 10 Parameter Code Hyperparameter Tuning Random Forest	75
Tabel 4. 11 Sebelum Tuning RF (Training)	79
Tabel 4. 12 Sebelum Tuning RF (Testing)	80
Tabel 4. 13 Hyperparameter Tuning RF (Training Set)	80
Tabel 4. 14 Hyperparameter Tuning RF (Testing Set)	81
Tabel 4. 15 Parameter Code Hyperparameter Tuning XGBoost	83
Tabel 4. 16 Sebelum Tuning XGBoost (Training)	86
Tabel 4. 17 Sebelum Tuning XGBoost (Testing)	87
Tabel 4. 18 Sesudah Tuning XGBoost (Training)	87
Tabel 4. 19 Sesudah Tuning XGBoost (Testing)	88
Tabel 4. 20 Sebelum Tuning LightGBM (Training)	92
Tabel 4. 21 Sebelum Tuning LightGBM (Testing)	92
Tabel 4. 22 Setelah Tuning LightGBM (Training)	93
Tabel 4. 23 Setelah Tuning LightGBM (Testing)	93
Tabel 4. 24 Parameter Code Hyperparameter Tuning CatBoost	95
Tabel 4. 25 Sebelum Tuning CatBoost (Training)	97

Tabel 4. 26 Sebelum Tuning CatBoost (Testing)	98
Tabel 4. 27 Sesudah Tuning CatBoost (Training)	98
Tabel 4. 28 Sesudah Tuning CatBoost (Testing)	99
Tabel 4. 29 Parameter Code Perbandingan Algoritma	102
Tabel 4. 30 Perbandingan Hasil Evaluasi Semua Algoritma	106
Tabel 4. 31 Penjelasan Fitur.....	110
Tabel 4. 32 Kolerasi Fitur Dengan Label.....	117



DAFTAR GAMBAR

Gambar 3. 1 Flowchart	30
Gambar 3. 2 Heatmap Korelasi	31
Gambar 3. 3 ROC Curve	42
Gambar 4. 1 Hasil SMOTETomek	62
Gambar 4. 2 Model Random Forest	65
Gambar 4. 3 Model XGBoost	68
Gambar 4. 4 Model LightGBM	71
Gambar 4. 5 Model CatBoost	74
Gambar 4. 6 Confusion Matrix RF Sebelum Tuning	82
Gambar 4. 7 Confusion Matrix RF Setelah Tuning	82
Gambar 4. 8 Confusion Matrix XGBoost Sebelum Tuning	88
Gambar 4. 9 Confusion Matrix XGBoost setelah Tuning	89
Gambar 4. 10 Confusion Matrix LightGBM Sebelum Tuning	94
Gambar 4. 11 Confusion Matrix LightGBM Setelah Tuning	94
Gambar 4. 12 Confusion Matrix CatBoost Sebelum Tuning	99
Gambar 4. 13 Confusion Matrix CatBoost Setelah Tuning	100
Gambar 4. 14 Perbandingan Algoritma	107
Gambar 4. 15 Deploy	108
Gambar 4. 16 Normal Traffic	109
Gambar 4. 17 Attack Traffic	109

INTISARI

Deteksi serangan jaringan merupakan aspek penting dalam menjaga keamanan sistem informasi, terutama ketika volume data yang masuk bersifat besar dan kompleks. Tantangan utama dalam proses ini adalah membangun model klasifikasi yang tidak hanya akurat, tetapi juga mampu menangani ketidakseimbangan kelas antara koneksi normal dan serangan. Untuk itu, penelitian ini membandingkan empat algoritma klasifikasi, yaitu Random Forest, XGBoost, LightGBM, dan CatBoost dalam mendeteksi serangan berbasis dataset UNSW-NB15. Penelitian dilakukan melalui tahapan eksplorasi data, preprocessing, penyeimbangan kelas menggunakan SMOTETomek, dan pelatihan model dengan tuning parameter melalui GridSearchCV. Evaluasi dilakukan berdasarkan metrik Accuracy, Precision, Recall, F1-Score, dan ROC AUC. Keempat algoritma menunjukkan kinerja tinggi dengan akurasi mencapai 90%, F1-score hingga 0,91, serta nilai precision dan recall yang seimbang. Namun, LightGBM mencatat ROC AUC tertinggi yaitu 0,987, dan menunjukkan performa paling optimal dan konsisten dibandingkan algoritma lainnya. LightGBM juga unggul dalam efisiensi pelatihan dan kecepatan inferensi, menjadikannya kandidat terbaik untuk diterapkan dalam sistem deteksi serangan di dunia nyata. Dengan hasil tersebut, penelitian ini merekomendasikan LightGBM sebagai algoritma yang paling tepat untuk diterapkan pada sistem deteksi intrusi jaringan, khususnya untuk data berskala besar dan kompleks. Ke depan, penelitian dapat dikembangkan dengan menguji integrasi model ke dalam sistem real-time dan menerapkan metode seleksi fitur untuk peningkatan efisiensi.

Kata kunci: Deteksi serangan, LightGBM, Klasifikasi, ROC AUC, *Machine Learning*

ABSTRACT

Intrusion detection is a crucial aspect of securing information systems, especially when dealing with large-scale and complex network traffic. One of the main challenges in this process is to develop a classification model that is not only accurate but also capable of handling class imbalance between normal and attack connections. This study compares the performance of four classification algorithms: Random Forest, XGBoost, LightGBM, and CatBoost for detecting network attacks using the UNSW-NB15 dataset. The research involved several stages, including data exploration, preprocessing, class balancing using SMOTETomek, and model training with hyperparameter tuning via GridSearchCV. The models were evaluated using Accuracy, Precision, Recall, F1-Score, and ROC AUC metrics. All four algorithms demonstrated strong performance, achieving 90% accuracy, F1-scores up to 0.91, and balanced precision and recall. However, LightGBM achieved the highest ROC AUC score of 0.987, indicating superior ability to distinguish between normal and malicious network traffic. LightGBM also offered faster training time and better efficiency, making it the most optimal and consistent model among those evaluated. Based on these results, LightGBM is recommended as the most suitable algorithm for implementing intrusion detection systems, particularly in scenarios involving large and complex datasets. Future research may include real-time system integration and the application of feature selection methods to further enhance model efficiency.

Keyword: Intrusion detection, LightGBM, Classification, ROC AUC, Machine Learning