

**ANALISIS PERBANDINGAN DAN INTERPRETASI SHAP VALUES  
TERHADAP MODEL *RANDOM FOREST*, *DECISION TREE*, DAN *XG-  
BOOST* DALAM DETEKSI INTRUSI JARINGAN**

**SKRIPSI**

**HALAMAN JUDUL**



**Di susun oleh:**

**Rafif Fuaddoshofha**

**20190140129**

**PROGRAM STUDI TEKNOLOGI INFORMASI  
FAKULTAS TEKNIK  
UNIVERSITAS MUHAMMADIYAH YOGYAKARTA**

**2024**

## HALAMAN PENYATAAN

Yang bertanda tangan di bawah ini:

**Nama** : Rafif Fuaddoshofha  
**NIM** : 20190140129  
**Program Studi** : Teknologi Informasi  
**Fakultas** : Teknik  
**Universitas** : Universitas Muhammadiyah Yogyakarta

Menyatakan dengan sebenar-benarnya bahwa tugas akhir ini yang berjudul "Analisis Perbandingan dan interpretasi shap values terhadap model Random Forest, Decision Tree, dan XG-Boost dalam deteksi intrusi jaringan". karya ini merupakan karya asli saya sendiri yang dibuat dengan bantuan dan bimbingan dari dosen pembimbing. Karya ini bukan hasil karya orang lain kecuali acuan atau kutipan yang telah disebutkan sumbernya. Karya ini belum pernah diajukan untuk memperoleh gelar akademik di Universitas Muhammadiyah Yogyakarta maupun institusi lainnya.

Yogyakarta, 08 April 2024

Yang Menyatakan,



A handwritten signature in black ink is written over a 10,000 Rupiah stamp. The stamp features the Garuda Pancasila emblem and the text '10000', 'METERAI TEMPEL', and the alphanumeric code 'A1042ALX105069207'.

**Rafif Fuaddoshofha**

## **MOTTO**

*“Memulai dengan penuh keyakinan, Menjalankan dengan penuh keikhlasan,  
Menyelesaikan dengan penuh kebahagiaan”*

*“jawaban dari sebuah keberhasilan adalah terus belajar dan tak kenal putus asa.”*

*“Aku tidak peduli atas keadaan susah dan senangku, karena aku tidak tahu manakah di  
antara keduanya itu yang lebih baik bagiku.” –Umar bin khattab*

## KATA PENGANTAR

Alhamdulillah Rabbil 'Alamin, Puji syukur kita panjatkan ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya yang senantiasa melimpahkan keberkahan sehingga penulis dapat menyelesaikan Tugas akhir dengan judul **“analisis perbandingan dan interpretasi Shap Value terhadap model (Random forest, decision tree, XG-Boost) dalam Deteksi Intrusi jaringan”**. Shalawat serta salam tak henti kita curahkan kepada junjungan kita Nabi Muhammad SAW, kemudian kepada para keluarga beliau, sahabat-sahabat beliau dan kepada pengikutnya sehingga kita bisa mendapatkan syafa'atnya hingga akhir zaman.

tugas akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana di Jurusan Teknologi Informasi di Universitas Muhammadiyah Yogyakarta. Dalam penyusunan tugas akhir ini penulis banyak mendapatkan bantuan serta bimbingan dari banyak pihak. Terutama karena adanya kehendak serta kemudahan yang diberikan oleh Allah'azza wa jalla penulis dapat menyelesaikan tugas akhir ini, kemudian pada kesempatan ini penulis ingin juga mengucapkan rasa Terimakasih yang sebesar-besarnya kepada :

1. Kepada Ayah saya Teguh utoyo dan ibunda saya Upidayani dan adik saya Daffa alfairoez Ramadhan, yang selalu memberikan dukungan serta motivasi dan Do'a restu sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Bapak Asroni S.T., M.Eng. selaku Dosen Pembimbing pertama yang telah sabar membimbing dan bersedia memberi masukan serta arahan kepada penulis dalam pembuatan skripsi hingga selesai.
3. Bapak Cahya Damarjati S.T., M.Eng. Ph.D. selaku Dosen Pembimbing kedua yang telah bersedia untuk membimbing dan memberi arahan dalam menyelesaikan skripsi ini.

4. Bapak Reza Giga Isnanda, S.T., M.Sc. selaku Kepala Program Studi Teknologi Informasi Universitas Muhammadiyah Yogyakarta.
5. Seluruh dosen, pengajar, staff Tata Usaha Prodi Teknologi Informasi Universitas Muhammadiyah Yogyakarta.
6. teman teman angkatan 2019 yang tidak bisa penulis sebutkan satu persatu namanya. Terimakasih sudah membersamai dan membantu penulis dalam masa perkuliahan.
7. Semua pihak yang telah membantu dalam penulisan dan penyelesaian skripsi yang tidak dapat disebutkan satu persatu.

Semoga Allah SWT. memberikan balasan yang berlipat ganda atas segala kebaikan yang telah dilakukan. Mungkin dalam Tugas Akhir banyak kesalahan yang telah penulis lakukan, untuk itu penulis ucapkan mohon maaf sebesar-besarnya, serta pembaca dapat memberi kritik dan saran yang membangun sehingga dapat membantu dalam penyusunan tugas akhir ini dapat lebih baik lagi.

Yogyakarta, 08 April 2024



Rafif Fuaddoshofha

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN I .....	ii
HALAMAN PENGESAHAN II .....	iii
HALAMAN PENYATAAN .....	iv
MOTTO.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI .....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiv
ABSTRACT .....	xv
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah .....	4
1.3. Batasan Masalah .....	4
1.4. Tujuan Penelitian .....	4
1.5. Manfaat Penelitian.....	5
1.6. Sistematika Penulisan.....	5
BAB II TINJUAN PUSTAKAN DAN LANDASAN TEORI .....	6
2.1. Tinjauan Pustaka.....	6
2.2. Landasan Teori.....	8
2.2.1 Intrusion Detection System .....	8
2.2.2 Machine Learning.....	9
2.2.3 Machine Learning Untuk Deteksi Intrusi .....	9
2.2.4 Dataset KDD Cup 99 .....	10
2.2.5 Preprocessing Data .....	10

2.2.6	Klasifikasi.....	10
2.2.7	Random Forest.....	11
2.2.8	Decision Tree.....	12
2.2.9	Extreme Gradient Boost (XG-boost).....	13
2.2.10	Evaluasi Matrix.....	13
2.2.11	SHAP Values (Shapley Additive Explanation).....	15
BAB III METODOLOGI PENELITIAN.....		17
3.1.	Pengambilan Dataset.....	18
3.2.	Pre Processing Data.....	20
3.2.1	Labeling Data.....	20
3.2.2	Pengkategorian Data atau Transformation data.....	20
3.2.3	Label Encoder dan feature mapping.....	21
3.2.4	Cleaning Data.....	21
3.2.5	Feature Selection.....	21
3.2.6	Split Data atau Pembagian Data.....	21
3.3.	Modeling.....	22
3.4.	Evaluasi Model.....	23
3.5.	Interpretasi Model SHAP values.....	24
BAB IV HASIL DAN PEMBAHASAN.....		25
4.1.	Pengambilan dataset.....	25
4.2.	Preprocessing Data.....	26
4.2.1	Labeling Data.....	26
4.2.2	Pengkategorian data atau Transformation data.....	27
4.2.3	Label Encoder Dan Feature Mapping.....	28
4.2.4	Cleaning Data.....	28
4.2.5	Feature Selection.....	29
4.2.6	Split Data atau Pembagian data.....	31
4.3.	Modeling Dan evaluasi Hasil.....	31
4.3.1	Random Forest.....	31

4.3.2	Decision Tree .....	33
4.3.3	XG-Boost.....	34
4.3.4	Perbandingan Model.....	36
4.4.	Interpretasi model Dengan SHAP Values .....	38
4.4.1	Interprtasi Model Random Forest.....	38
4.4.2	Interpretasi model Decision tree.....	46
4.4.3	Interpretasi Model XG-Boost .....	53
BAB V KESIMPULAN DAN SARAN.....		60
5.1.	Kesimpulan.....	60
5.2.	Saran .....	61
DAFTAR PUSTAKA .....		62



## DAFTAR TABEL

Table 2.2.10. 1 Confusion Matrix.....	14
Tabel 3. 1 Deskripsi Fitur.....	18
Tabel 3. 2 Deskripsi kategori Serangan.....	20
Tabel 3. 3 parameter yang di gunakan pada setiap model algoritma.....	22
Tabel 4.2. 1 Nama Fiur .....	26
Tabel 4.2.2. 1 Jenis Sernagan yang dalam dataset kdd cup 99 .....	27
Tabel 4.2.2. 2 Pengkategorian Jenis Serangan.....	27
Table 4.2.5. 1 Fitur yang dihapus dari hasil correlation .....	29
Table 4.2.5. 2 Fitur yang digunakan .....	30
Table 4.3. 1 Performa Random Forest.....	31
Table 4.3. 2 Evaluasi perhitungan performa Matrix Random Forest .....	32
Table 4.3. 3 Performa Decision Tree .....	33
Table 4.3. 4 Evaluasi perhitungan performa Matrix Decision Tree.....	33
Table 4.3. 5 Peforma Xg-Boost .....	34
Table 4.3. 6 Evaluasi perhitungan performa Matrix Decision Tree.....	35
Tabel 4.4. 1 Penjelasan warna Batang .....	39

## DAFTAR GAMBAR

Gambar 2. 1 Flowchart Random Forest .....	11
Gambar 2. 2 Flow chart Decision Tree .....	12
Gambar 2. 3 gambaran cara kerja dari shap values .....	16
Gamb 3. 1 Alur Penelitian .....	17
Gambar 4. 1 sampel data .....	25
Gambar 4.3. 1 Confusion Matrix Random Forest .....	32
Gambar 4.3. 2 Confusion Matrix Decision Tree .....	34
Gambar 4.3. 3 Confusion Matrix XG-Boost .....	35
Gambar 4.3. 4 Waktu Pengujian.....	36
Gambar 4.3. 5 Accuracy Model.....	36
Gambar 4.4.1. 1 Plot Summary Random Forest.....	38
Gambar 4.4.1. 2 Kelas DOS (0).....	40
Gambar 4.4.1. 3 Kelas Normal (1) .....	41
Gambar 4.4.1. 4 Probe (2) .....	42
Gambar 4.4.1. 5 R2L (3).....	43
Gambar 4.4.1. 6 U2R (4).....	44
Gambar 4.4.1. 7 Plot Force random forest .....	45
Gambar 4.4.2. 1 Plot Summary Decision Tree .....	46
Gambar 4.4.2. 2 Dos(0) .....	47
Gambar 4.4.2. 3 Normal(1) .....	48
Gambar 4.4.2. 4 Probe(2) .....	49
Gambar 4.4.2. 5 R2l(3).....	50
Gambar 4.4.2. 6 U2R(4).....	51
Gambar 4.4.2. 7 Plot Decision decision tree .....	52
Gambar 4.4.3. 1 Summary Plot XG-Boost.....	53
Gambar 4.4.3. 2 DOS(0).....	54
Gambar 4.4.3. 3 Normal (1) .....	55
Gambar 4.4.3. 4 Probe(2) .....	56

Gambar 4.4.3. 5 R2L (3).....	57
Gambar 4.4.3. 6 U2R(4).....	58
Gambar 4.4.3. 7 Decision Plot XG-Boost .....	59