

**ANALISIS PERBANDINGAN ALGORITMA *LIGHTGBM* DAN
RANDOM FOREST DALAM MENDETEKSI *PHISHING*
BERBASIS *URL***

TUGAS AKHIR



NANDA ALIF BAKRI

2101020120

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2026**

**ANALISIS PERBANDINGAN ALGORITMA *LIGHTGBM* DAN
RANDOM FOREST DALAM MENDETEKSI *PHISHING*
BERBASIS URL**

Tugas Akhir

Skema Skripsi

Tugas Akhir diajukan sebagai salah satu persyaratan memperoleh gelar Sarjana

Teknik pada Program Studi Teknik Informatika



NANDA ALIF BAKRI

2101020120

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2026**

PERNYATAAN MENGENAI TUGAS AKHIR DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tugas akhir berjudul “Analisis Perbandingan Algoritma *LightGBM* dan *Random Forest* dalam Mendeteksi *Phishing* Berbasis *URL*” adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan melanggar peraturan yang berlaku dalam karya tulis dan hak intelektual, maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Maritim Raja Ali Haji dan menerima sanksi lainnya sesuai dengan peraturan yang berlaku. Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Maritim Raja Ali Haji.

Tanjungpinang, Juli 2026



Nanda Alif Bakri
NIM 2101020120



© Hak Cipta Milik Universitas Maritim Raja Ali Haji, Tahun 2026

Hak Cipta Dilindungi Undang- Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah; dan pengutipan tersebut tidak merugikan kepentingan Universitas Maritim Raja Ali Haji.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Universitas Maritim Raja Ali Haji.

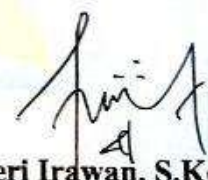


HALAMAN PERSETUJUAN

Judul : Analisis Perbandingan Algoritma *LightGBM* dan
Random Forest dalam Mendeteksi *Phishing* Berbasis
URL
Nama : Nanda Alif Bakri
NIM : 2101020120
Program Studi : Teknik Informatika
Tanggal Persetujuan : 10 Juni 2026

Disetujui oleh,
Komisi Pembimbing


Tekad Matulatan, S.Sos., S.Kom.,
M.Info.Tech
(Ketua)


Feri Irawan, S.Kom., M.Kom
(Anggota)

HALAMAN PENGESAHAN

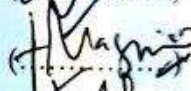
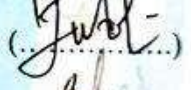
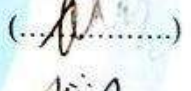
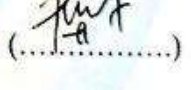
Judul : Analisis Perbandingan Algoritma *LightGBM* dan *Random Forest* dalam Mendeteksi *Phishing* Berbasis *URL*

Nama : Nanda Alif Bakri

NIM : 2101020120

Telah berhasil dipertahankan di hadapan Komisi Penguji sebagai bagian persyaratan yang diperlukan dalam memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Jurusan Teknik Elektro dan Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

KOMISI PENGUJI

Ketua	: Ferdi Chahyadi, S.Kom., M.Cs.	()
Anggota I	: Fortia Magfira, M.Kom.	()
Anggota II	: Muhamad Fadli, M.Kom.	()
Anggota III	: Tekad Matulatan, S.Sos., S.Kom., M.Info.Tech.	()
Anggota IV	: Feri Irawan, S.Kom., M.Kom.	()

Mengetahui,

Ketua Jurusan Teknik Elektro dan Informatika


Hollanda Arief Kusuma, S.IK., M.Si

NIP. 9904014019031016

Tanggal Ujian: 24 Juni 2026

Tanggal Lulus:

PRAKATA

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Perbandingan Algoritma *LightGBM* dan *Random Forest* dalam Mendeteksi *Phishing* Berbasis *URL*”. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) pada Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa terdapat banyak pihak yang telah memberikan bimbingan, bantuan, dukungan, dan motivasi, baik secara langsung maupun tidak langsung. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Martaleli Bettiza, S.Si., M.Sc., selaku Dekan Fakultas Teknik dan Teknologi Kemaritiman Universitas Maritim Raja Ali Haji.
2. Bapak Hollanda Arief Kusuma, S.T., M.Si., selaku Ketua Jurusan Teknik Elektro dan Informatika.
3. Bapak Tekad Matulatan, S.Sos., S.Kom., M.Info.Tech., selaku Dosen Pembimbing I yang telah meluangkan waktu untuk memberikan arahan, bimbingan, masukan, dan ilmu selama proses penyusunan skripsi ini.
4. Bapak Feri Irawan, S.Kom., M.Kom., selaku Dosen Pembimbing II yang telah memberikan saran, motivasi, dan bimbingan hingga skripsi ini dapat diselesaikan.
5. Seluruh dosen Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji yang telah memberikan ilmu, pengalaman, dan bimbingan selama masa perkuliahan.
6. Seluruh staf administrasi Fakultas Teknik dan Teknologi Kemaritiman Universitas Maritim Raja Ali Haji yang telah membantu kelancaran proses akademik dan penyusunan skripsi.
7. Orang tua tercinta yang selalu memberikan kasih sayang, doa, dukungan, semangat, dan motivasi tanpa henti kepada penulis.
8. Seluruh anggota keluarga yang senantiasa memberikan perhatian, bantuan, dan doa selama proses perkuliahan hingga penyelesaian skripsi.

9. Teman-teman seperjuangan di Program Studi Teknik Informatika yang telah memberikan bantuan, semangat, dan kebersamaan selama masa perkuliahan.
10. Bang Luthfi Halimawan dan Setiawan Ade, selaku *streamer* dan kreator konten yang melalui siaran serta kontennya telah memberikan hiburan, semangat, dan motivasi kepada penulis selama proses perkuliahan dan penyusunan skripsi.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan karena keterbatasan pengetahuan dan kemampuan yang dimiliki. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun sebagai bahan perbaikan pada masa yang akan datang. Penulis berharap skripsi ini dapat memberikan manfaat bagi pembaca dan turut berkontribusi terhadap pengembangan ilmu pengetahuan, khususnya dalam bidang keamanan siber, *machine learning*, dan deteksi URL *phishing*.

Tanjungpinang, Juli 2026



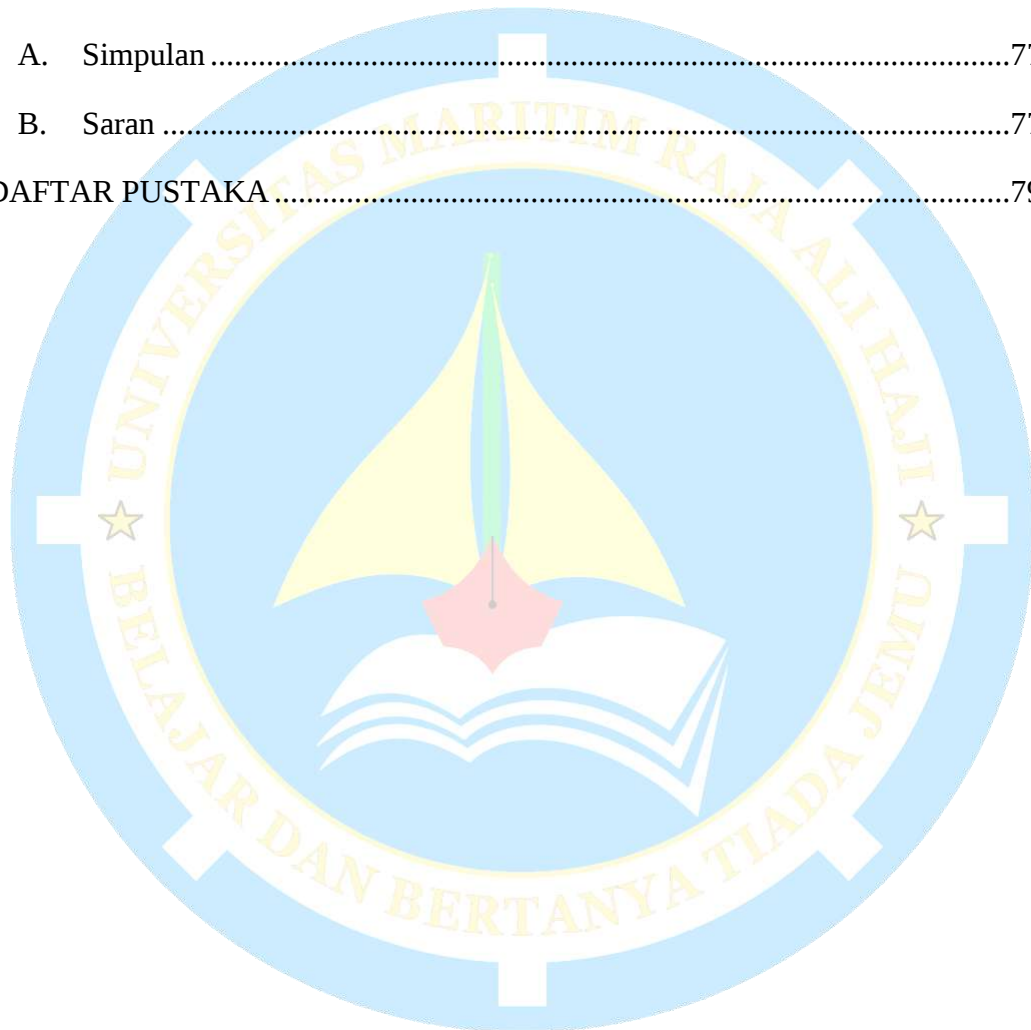
Nanda Alif Bakri

DAFTAR ISI

HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN	v
ABSTRAK.....	vi
ABSTRACT.....	vii
PRAKATA.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN.....	1
A. Latar Belakang.....	1
B. Perumusan Masalah.....	2
C. Batasan Penelitian.....	2
D. Tujuan Penelitian.....	3
E. Manfaat Penelitian.....	3
BAB II KAJIAN LITERATUR.....	4
A. Tinjauan Pustaka.....	4
B. Landasan Teori.....	6
1. Keamanan Siber.....	6
2. <i>Phishing</i>	8
3. <i>Legitimate URL</i>	9
4. Deteksi <i>Phishing</i>	10
5. <i>Machine Phishing</i>	11
6. <i>Ensemble Phishing</i>	11
7. <i>Random Forest</i>	13

8.	<i>LightGBM</i>	16
9.	<i>Hyperparameter</i>	19
10.	<i>Confusion matrix</i>	20
11.	<i>Python</i>	21
12.	<i>Google Colab</i>	22
BAB III METODE PENELITIAN		23
A.	Waktu dan Tempat Penelitian	23
B.	Bahan dan Alat	23
C.	Prosedur Pelaksanaan Penelitian	23
D.	Studi Pustaka	25
E.	Pengumpulan Data	25
F.	Analisis Dan Perancangan	27
1.	Analisis Data	27
2.	★ Perancangan Algoritma <i>Random Forest</i>	33
3.	Perancangan Algoritma <i>LightGBM</i>	43
G.	Perancangan Sistem	53
1.	<i>Use Case Diagram</i>	54
2.	<i>Activity Diagram</i>	55
3.	Perancangan <i>User Interface (UI)</i>	56
BAB IV HASIL DAN PEMBAHASAN		57
A.	Pelatihan dan Evaluasi Model	57
1.	Pelatihan dan Evaluasi Algoritma <i>Random Forest</i>	58
2.	Pelatihan dan Evaluasi Algoritma <i>LightGBM</i>	62
3.	Perbandingan Kinerja Model pada Variasi Pembagian Data	67
B.	Pembahasan	68
C.	Hasil Perancangan Aplikasi <i>Web</i>	70

D. Pengujian Model pada Data Baru melalui Aplikasi <i>Web</i>	71
1. Pengujian <i>URL</i> Pendek.....	72
2. Pengujian <i>URL</i> Panjang.....	73
3. Pengujian <i>URL</i> dengan karakter <i>Non-Latin</i>	73
4. Pengujian <i>URL</i> dengan Probabilitas Mendekati Ambang Klasifikasi.....	75
BAB V SIMPULAN DAN SARAN.....	77
A. Simpulan	77
B. Saran	77
DAFTAR PUSTAKA	79



DAFTAR TABEL

Tabel 1. <i>Confusion matrix</i>	21
Tabel 2. Bahan yang digunakan.....	23
Tabel 3. Alat yang digunakan	23
Tabel 4. Sampel Hasil Ekstraksi Fitur <i>URL</i>	33
Tabel 5. Dataset <i>URL</i>	37
Tabel 6. <i>Node</i> kiri	38
Tabel 7. <i>Node</i> kanan	39
Tabel 8. Hasil pohon pertama	40
Tabel 9. Hasil prediksi <i>Random Forest legitimate</i>	42
Tabel 10. Hasil prediksi <i>Random Forest phishing</i>	43
Tabel 11. Dataset <i>URL</i>	47
Tabel 12. Hasil prediksi Awal	48
Tabel 13. Hasil <i>residual</i> pertama	48
Tabel 14. Hasil prediksi setelah pohon pertama	50
Tabel 15. Hasil <i>residual</i> kedua	50
Tabel 16. Hasil prediksi setelah pohon kedua	51
Tabel 17. Hasil prediksi setelah pohon ketiga	53
Tabel 18. Hasil prediksi akhir <i>LightGBM</i>	53
Tabel 19. Hasil Evaluasi Model <i>Random Forest</i>	59
Tabel 20. Hasil Evaluasi Model <i>LightGBM</i>	64
Tabel 21. Jumlah Data pada Setiap Skenario Pembagian.....	67
Tabel 22. Perbandingan Model <i>Radom Forest</i>	67
Tabel 23. Perbandingan Model <i>LightGBM</i>	67
Tabel 24. Perbandingan Performa Model	68
Tabel 25. Perbandingan Efisiensi Komputasi	68
Tabel 26. Hasil Pengujian <i>URL</i> Pendek berlabel <i>Phishing</i>	72
Tabel 27. Hasil Pengujian <i>URL</i> Panjang Berlabel <i>Phishing</i>	73
Tabel 28. Hasil Pengujian <i>URL</i> dengan Karakter <i>Non-Latin</i>	74
Tabel 29. Hasil Pengujian <i>URL</i> Mendekati Ambang Klasifikasi	76

DAFTAR GAMBAR

Gambar 1. <i>CIA Triad</i> (Sumber: [38])	7
Gambar 2. Struktur Algoritma <i>Random Forest</i> (Sumber: [39])	14
Gambar 3. Arsitektur Algoritma <i>LightGBM</i> (Sumber: [40])	17
Gambar 4. Prosedur Penelitian	24
Gambar 5. Sebaran Dataset <i>Phishing</i> dan <i>Legitimate</i>	26
Gambar 6. Dataset yang digunakan	27
Gambar 7. <i>Flowchart</i> Algoritma <i>Random Forest</i>	34
Gambar 8. <i>Flowchart</i> Algoritma <i>LightGBM</i>	44
Gambar 9. <i>Use Case Diagram</i>	54
Gambar 10. <i>Activity Diagram</i>	55
Gambar 11. Halaman Deteksi	56
Gambar 12. Grafik <i>Tuning Hyperparameter Random Forest</i>	59
Gambar 13. <i>Confusion matrix Random Forest</i>	60
Gambar 14. Grafik <i>Training dan Validation Loss Model Random Forest</i>	61
Gambar 15. Grafik <i>Tuning Hyperparameter LightGBM</i>	63
Gambar 16. <i>Confusion matrix LightGBM</i>	65
Gambar 17. Grafik <i>Training dan Validation loss Model LightGBM</i>	66
Gambar 18. Halaman Home dengan Hasil Deteksi <i>Legitimate</i>	70
Gambar 19. Halaman Home dengan Hasil Deteksi <i>Phishing</i>	71
Gambar 20. Hasil pengujian <i>URL</i> dengan Karakter <i>Non-Latin</i>	74
Gambar 21. Hasil Pengujian <i>URL</i> dengan Probabilitas Mendekati Ambang Klasifikasi	75