DETEKSI WEBSITE PHISHING BERBASIS KARAKTERISTIK URL MENGGUNAKAN PENDEKATAN RANDOM FOREST DAN DECISION TREE

SKRIPSI

LOUIS GAUDIAMO HALIM
00000062782

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026

DETEKSI WEBSITE PHISHING BERBASIS KARAKTERISTIK URL MENGGUNAKAN PENDEKATAN RANDOM FOREST DAN DECISION TREE

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

LOUIS GAUDIAMO HALIM
00000062782

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2026

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama	:	Louis Gaudiamo Halim
Nomor Induk Mahasiswa	:	00000062782
Program Studi	:	Informatika

Skripsi dengan judul:

Deteksi Website Phishing Berbasis Karakteristik URL Menggunakan Pendekatan Random Forest dan Decision Tree

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 08 Januari 2026

(Louis Gaudiamo Halim)

HALAMAN PENGESAHAN

Skripsi dengan judul

DETEKSI WEBSITE PHISHING BERBASIS KARAKTERISTIK URL MENGGUNAKAN PENDEKATAN RANDOM FOREST DAN DECISION TREE

oleh

Nama	:	Louis Gaudiamo Halim
NIM	:	00000062782
Program Studi	:	Informatika
Fakultas	:	Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 15 Januari 2026

Pukul 15.00 s/s 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

(Wirawan Istiono, S.Kom., M.Kom)

NIDN: 0313048304

Penguji

23/1 2026.

(Fenina Adline Twince Tobing, S.Kom., M.Kom)

NIDN: 0406058802

Pembimbing

26/1 2026

(SY.Yuliani,S.Kom.,MT.,Ph.D.,CEH)

NIDN: 0411037904

Ketua Program Studi Informatika,

(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini:

Nama	:	Louis Gaudiamo Halim
NIM	:	00000062782
Program Studi	:	Informatika
Jenjang	:	S1
Judul Karya Ilmiah	:	Deteksi Website Phishing Berbasis Karakteristik URL Menggunakan Pendekatan Random Forest dan Decision Tree

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

☑ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.

☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.

☐ Lainnya, pilih salah satu:

- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
- Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 08 Januari 2026

Yang menyatakan

Louis Gaudiamo Halim

HALAMAN PERSEMBAHAN / MOTTO

”Kejahatan siber merupakan ancaman terbesar bagi setiap perusahaan di dunia.”

Eugene Kaspersky

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan karunia-Nya, sehingga laporan Tugas Akhir dengan judul ”Deteksi Website Phishing Berbasis Karakteristik URL Menggunakan Pendekatan Random Forest dan Decision Tree” dapat terselesaikan dengan baik. Laporan ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana Komputer di Universitas Multimedia Nusantara. Penulis menyadari bahwa kelancaran penyusunan laporan ini tidak lepas dari bimbingan dan dukungan berbagai pihak.
Mengucapkan terima kasih

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu SY.Yuliani.,S.Kom.,MT.,Ph.D.,CEH, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Teman-teman yang telah membantu dan mendukung saya selama penulisan laporan skripsi ini.

Semoga laporan Tugas Akhir ini dapat memberikan manfaat dan kontribusi positif bagi pembaca serta pengembangan ilmu pengetahuan.

Tangerang, 08 Januari 2026

Louis Gaudiamo Halim

DETEKSI WEBSITE PHISHING BERBASIS KARAKTERISTIK URL MENGGUNAKAN PENDEKATAN RANDOM FOREST DAN DECISION TREE

Louis Gaudiamo Halim

ABSTRAK

Serangan *phishing* merupakan teknik rekayasa sosial yang menjadi salah satu ancaman keamanan siber terbesar yang bertujuan mencuri informasi sensitif dengan memalsukan situs web agar terlihat asli. Tantangan utama sistem deteksi saat ini adalah bergantung pada analisis isi halaman (*content-based*) sering kali membebani sistem karena membutuhkan daya komputasi tinggi dan waktu proses yang lama. Untuk mengatasi kendala tersebut, penelitian ini menawarkan solusi yang lebih ringan dan cepat, yaitu sistem pendeteksi *phishing* hanya dengan menganalisis karakteristik URL (*lexical-based features*) dan reputasi domain, tanpa perlu mengunduh konten halaman. Dengan memanfaatkan dataset publik dari Kaggle yang mencakup 11.430 sampel URL yang terdiri dari 89 kolom fitur, studi ini menguji ketangguhan algoritma *Random Forest* melawan model dasar *Decision Tree*. Melalui serangkaian pengujian fitur leksikal dan optimasi *hyperparameter tuning* untuk hasil maksimal. Hasil eksperimen membuktikan keunggulan *Random Forest*, yang berhasil mencatat akurasi (*accuracy*) hingga 95.51%, dengan presisi (*precision*) 95.51%, *recall* 95.43%, dan *F1-score* 95.47%, hasil yang lebih stabil dibandingkan *Decision Tree* di angka 92.52%. Temuan ini menegaskan bahwa kombinasi fitur sederhana seperti umur domain dan status indeks pencarian sudah cukup untuk membangun pertahanan anti-*phishing* yang cepat dan akurat.

Kata kunci: Karakteriktik URL, Keamanan Siber, *Machine Learning*, *Phishing*, *Random Forest*

PHISHING WEBSITE DETECTION BASED ON URL FEATURES USING RANDOM FOREST AND DECISION TREE

Louis Gaudiamo Halim

ABSTRACT

Phishing attacks are social engineering techniques representing one of the greatest cybersecurity threats, aiming to steal sensitive information by falsifying websites to appear authentic. The main challenge of current detection systems lies in their reliance on content-based analysis, which often burdens the system due to high computational power requirements and long processing times. To address these constraints, this study offers a lighter and faster solution include a phishing detection system that analyzes only URL characteristics (lexical-based features) and domain reputation, eliminating the need to download page content. Utilizing a public dataset from Kaggle comprising 11,430 URL samples and 89 feature columns, this study evaluates the robustness of the Random Forest algorithm against a baseline Decision Tree model. The process involves a series of lexical feature tests and hyperparameter tuning optimization to achieve maximum results. Experimental results demonstrate the superiority of Random Forest, achieving an accuracy of 95.51%, with a precision of 95.51%, recall of 95.43%, and F1-score of 95.47%, reflecting significantly higher stability compared to the Decision Tree at 92.52%. These findings confirm that a combination of simple features, such as domain age and search index status, is sufficient to construct a fast and accurate anti-phishing defense.

Keywords: *Cybersecurity*, *Machine Learning*, *Phishing*, *Random Forest*, URL Characteristics

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR KODE

DAFTAR RUMUS

DAFTAR LAMPIRAN