

**SISTEM DETEKSI *WEBSITE PHISHING* DENGAN
ALGORITMA KLASIFIKASI *XGBOOST* DAN
RANDOM FOREST BERBASIS WEB**

SKRIPSI

Oleh:

Devi Putri Ningsih

202110715177



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA**

2025

LEMBAR PERSETUJUAN PEMBIMBING

(DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Skripsi : Sistem Deteksi *Website Phishing* Dengan
Algoritma Klasifikasi *XGboost* Dan *Random*
Forest Berbasis Web

Nama Mahasiswa : Devi Putri Ningsih

Nomor Pokok Mahasiswa : 202110715177

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 28 Juli 2025

Jakarta, 04 Agustus 2025

MENYETUJUI,

Pembimbing I

Pembimbing II

Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.
NIDN. 0322108201

Nurfiyah, S.T., M.Kom.
NIDN. 0312079403

Ketua Program Studi

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya
2025

LEMBAR PENGESAHAN

Judul Tugas Akhir : Sistem Deteksi *Website Phishing* Dengan
Algoritma Klasifikasi *XGboost* Dan *Random*
Forest Berbasis Web

Nama Mahasiswa : Devi Putri Ningsih

Nomor Pokok Mahasiswa : 202110715177

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 28 Juli 2025

Jakarta, 04 Agustus 2025

MENGESAHKAN,

Ketua Tim Penguji : Prof. Dr. Mujiono Sadikin, S.T., M.T.
NIDN. 0406127002

Penguji I : Khairunnisa Fadhilla Ramdhania, S.Si., M.Si.
NIDN. 0328039201

Penguji II : Rakhmat Purnomo, S.Pd, S.Kom., M.Kom.
NIDN. 0322108201

MENGETAHUI,

Ketua
Program Studi Informatika

Ahmad Fathurrozi, S.E., M.M.S.I
NIP. 2012486

Dekan
Fakultas Ilmu Komputer

Dr. Robertus Suraji, M.A.
NIP. 1904406

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini:

Nama : Devi Putri Ningsih

NPM : 202110715177

Program Studi : Informatika

Fakultas : Ilmu Komputer

Judul Tugas Akhir : Sistem Deteksi *Website Phishing* Dengan Algoritma Klasifikasi *XGboost* Dan *Random Forest* Berbasis Web

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 04 Agustus 2025



Devi Putri Ningsih

ABSTRAK

Devi Putri Ningsih, 202110715177. Sistem Deteksi Website Phishing Dengan Algoritma Klasifikasi *XGboost* Dan *Random Forest* Berbasis Web.

Perkembangan teknologi informasi telah mempermudah akses terhadap berbagai layanan daring, namun juga meningkatkan risiko kejahatan siber, salah satunya adalah *phishing*. *Phishing* merupakan upaya penipuan dengan menyamar sebagai situs terpercaya guna memperoleh data sensitif dari pengguna. Penelitian ini bertujuan untuk mengembangkan sistem deteksi website phishing berbasis web menggunakan algoritma klasifikasi *random forest* dan *XGBoost*. Sistem dirancang agar dapat menerima input URL secara manual, kemudian mengekstraksi fitur-fitur penting dari URL untuk dilakukan klasifikasi. Model dikembangkan menggunakan dataset dari URL, dengan pembagian data latih dan uji sebesar 80:20. Evaluasi performa dilakukan dengan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa *XGBoost* memiliki performa lebih unggul dengan akurasi 98,7%, dibandingkan *random forest* yang memiliki akurasi 96,4%. Sistem ini juga telah berhasil diuji secara fungsional dan mampu mendeteksi URL *phishing* dengan baik serta menampilkan hasil secara *real-time*. Penelitian ini menunjukkan bahwa *XGBoost* lebih layak digunakan dalam sistem deteksi phishing berbasis web, dan aplikasi yang dikembangkan dapat menjadi alat bantu bagi pengguna umum untuk mengidentifikasi ancaman phishing dengan lebih cepat dan akurat.

Kata kunci: *Phishing*, Deteksi Website, *Machine Learning*, *XGBoost*, *Random Forest*, klasifikasi, sistem berbasis web

ABSTRACT

Devi Putri Ningsih, 202110715177. Phishing Website Detection System With XGboost Classification Algorithm And Web-Based Random Forest.

The development of information technology has made it easier to access various online services, but it has also increased the risk of cybercrime, one of which is phishing. Phishing is an attempt to deceive by disguising itself as a trusted site in order to obtain sensitive data from users. This study aims to develop a web-based phishing website detection system using the Random Forest and XGBoost classification algorithms. The system is designed to be able to accept URL input, then extract important features from the URL for classification. The model was developed using a dataset from URL, with a training and test data division of 80:20. Performance evaluation was carried out using accuracy, precision, recall, and F1-score metrics. The test results showed that XGBoost had superior performance with an accuracy of 98.7%, compared to Random Forest which had an accuracy of 96.4%. This system has also been successfully tested functionally and is able to detect phishing URLs well and display results in real-time. This study shows that XGBoost is more feasible to be used in a web-based phishing detection system, and the developed application can be a tool for general users to identify phishing threats more quickly and accurately.

Keywords: *Phishing, Website Detection, Machine Learning, XGBoost, Random Forest, classification, web-based system*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini:

Nama : Devi Putri Ningsih
NPM : 202110715177
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif** (*Non-Exclusive Royalty-Free Right*), atas karya ilmiah saya yang berjudul :

Sistem Deteksi *Website Phishing* Dengan Algoritma Klasifikasi *XGBoost* Dan *Random Forest* Berbasis Web

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi

Pada tanggal : 04 Agustus 2025


3C0CCAMX440614654
Devi Putri Ningsih

KATA PENGANTAR

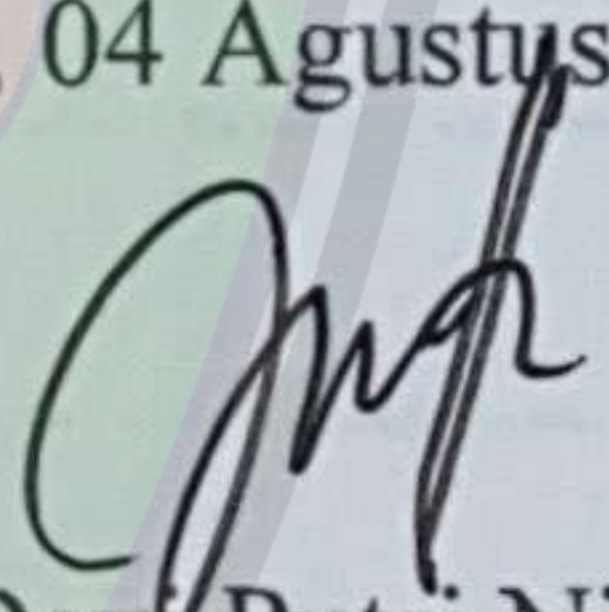
Segala puji bagi Allah swt, yang telah melimpahkan rahmat dan hidayah nya sehingga saya dapat menyelesaikan Tugas Akhir dengan semaksimal Mungkin demi kesempurnaan Penelitian Skripsi Yang berjudul “**Sistem Deteksi Website Phishing Dengan Algoritma Klasifikasi Xgboost Dan Random Forest Berbasis Web**” yang saya teliti dari dataset *website Phishing* yang saya cari dari *URL*. Pada penelitian ini saya ingin mengucapkan rasa terimakasih saya kepada:

1. Bapak Irjen Pol (Purn) Prof. Dr. Drs. Bambang Karsono, S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A. Selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. Selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Bapak Rakhmat Purnomo, S.Pd., S.Kom., M.Kom. Selaku Dosen Pembimbing I Skripsi Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
5. Ibu Nurfiyah, S.T., M.Kom. Selaku Dosen Pembimbing II dalam penyusunan laporan skripsi pada Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
6. Bapak Wowon Priatna, S.T., M.Ti. Selaku Dosen Pembimbing Akademi Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhyangkara Jakarta Raya.

7. Seluruh Dosen Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya yang telah Memberikan Ilmu dan Bimbingannya Selama Masa Perkuliahan.
8. Kepada Kedua Orang Tua Saya yang telah Memberikan dukungan Moral, Material serta Motivasi untuk selalu Kuat dan tetap Berusah agar semuanya terlaksana dengan baik.

Penelitian Skripsi yang saya susun ini, mungkin masih terdapat banyak kekurangan dan masih jauh dari sempurna, oleh karena itu saya dengan hati terbuka menerima masukan baik berupa kritik, maupun saran-saran yang dapat membangun untuk memperbaiki kekurangan yang ada. Semoga penelitian ini bermanfaat bagi saya pribadi maupun bagi orang lain yang membacannya.

Jakarta, 04 Agustus 2025


Devi Putri Ningsih

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	2
1.3 Rumusan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Batasan Masalah.....	4
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terdahulu	7

2.2	<i>Phishing</i>	11
2.3	<i>Website Phishing</i>	11
2.4	Sistem Deteksi <i>Phishing</i> Berbasis Web.....	16
2.5	Deteksi <i>Phishing</i> Berbasis <i>Machine learning</i>	17
2.6	Klasifikasi.....	18
2.7	<i>XGBoost (Extreme Gradient Boosting)</i>	20
2.8	<i>Random Forest</i>	21
2.9	PHP (<i>Hypertext preprocessor</i>).....	24
2.10	<i>Hyper Text Markup Language (HTML)</i>	24
2.11	<i>Web</i>	25
2.12	<i>XAMPP</i>	26
2.13	<i>Unified Modeling Language (UML)</i>	27
2.13.1	<i>Use Case Diagram</i>	28
2.13.2	<i>Activity Diagram</i>	29
2.13.3	<i>Sequence Diagram</i>	31
2.13.4	<i>Class Diagram</i>	32
2.14	<i>Blackbox Testing</i>	33
BAB III METODOLOGI PENELITIAN		35
3.1	Jenis Dan Pendekatan Penelitian	35
3.2	Kerangka Penelitian.....	35
3.3	Arsitektur Sistem.....	36
3.4	Teknik Pengumpulan Data	37
3.5	Metode Analisis Data	37

3.6	Pengujian Sistem	38
3.7	Analisis Kebutuhan Sistem (Hardware dan Software).....	38
BAB IV HASIL DAN PEMBAHASAN.....		40
4.1	Tahapan Algoritma <i>XGBoost</i> Dan <i>Random Forest</i>	40
4.2	Gambaran Umum Sistem	42
4.3	Perancangan Sistem.....	42
4.3.1	<i>Use Case Diagram</i>	43
4.3.2	<i>Activity Diagram</i> Sistem.....	44
4.4	Hasil Implementasi Sistem.....	46
4.4.1	Antarmuka Pengguna (<i>Frontend</i>).....	46
4.4.2	<i>Backend</i> dan Proses Klasifikasi	49
4.4.3	Integrasi Model <i>Machine Learning</i>	49
4.5	Hasil Pengujian Sistem.....	50
4.6.1	Pengujian Fungsionalitas Sistem.....	50
4.6.2	Metode Evaluasi.....	51
4.6.3	Hasil Evaluasi Model	51
4.6	Pembahasan.....	52
BAB V PENUTUP		56
5.1	Kesimpulan.....	56
5.2	Saran.....	57
DAFTAR PUSTAKA.....		58
LAMPIRAN.....		64

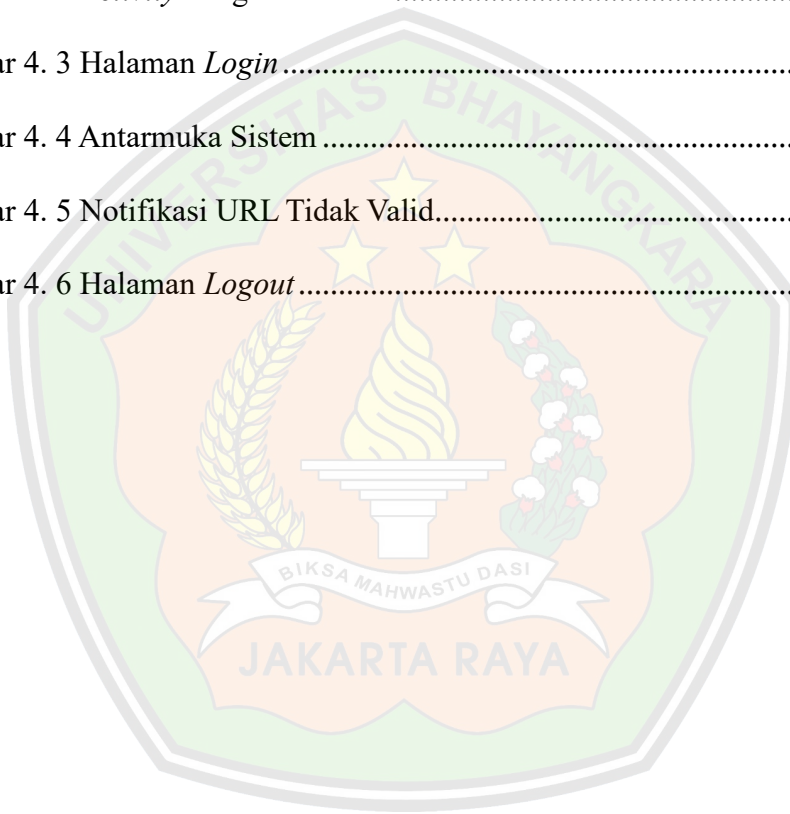
DAFTAR TABEL

Tabel 2. 2 Simbol-Simbol <i>Use Case Diagram</i>	28
Tabel 2. 3 Simbol-simbol <i>Activity Diagram</i>	29
Tabel 2. 4 Simbol-simbol <i>Sequence Diagram</i>	31
Tabel 2. 5 Simbol-simbol <i>Class Diagram</i>	32
Tabel 4. 1 Ekstraksi Fitur URL	40
Tabel 4. 2 <i>Output</i> Simulasi Probabilitas.....	41
Tabel 4. 5 Pengujian Fungsionalitas Sistem.....	50
Tabel 4. 6 Hasil Evaluasi Model	52



DAFTAR GAMBAR

Gambar 2. 1 <i>Interface</i> XAMPP.....	27
Gambar 3. 1 Kerangka Penelitian	36
Gambar 4. 1 <i>Use Case</i> Diagram Sistem.....	43
Gambar 4. 2 <i>Activity</i> Diagram Sistem.....	45
Gambar 4. 3 Halaman <i>Login</i>	47
Gambar 4. 4 Antarmuka Sistem	48
Gambar 4. 5 Notifikasi URL Tidak Valid.....	48
Gambar 4. 6 Halaman <i>Logout</i>	49



DAFTAR LAMPIRAN

	Hal
Lampiran 1. 1 Hasil Plagiarisme	65
Lampiran 1. 2 Biodata Mahasiswa	66
Lampiran 1. 3 Rekomendasi Pembimbing untuk Skripsi	67
Lampiran 1. 4 Lembar Bimbingan Skripsi Dengan Pembimbing I	69
Lampiran 1. 5 Lembar Bimbingan Skripsi Dengan Pembimbing II	70

