DETEKSI PHISHING MENGGUNAKAN INFORMATION GAIN DAN RANDOM FOREST

SKRIPSI

Oleh:

ALFAJAR ALVIN PERMANA TAMBUNAN
221112970
DWI AMANDA MONICA SIMANJUNTAK
221113275
RIYANDO RAJAGUKGUK
221112489

PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026

PHISHING DETECTION USING INFORMATION GAIN AND RANDOM FOREST

UNDERGRADUATE THESIS

By:

ALFAJAR ALVIN PERMANA TAMBUNAN
221112970
DWI AMANDA MONICA SIMANJUNTAK
221113275
RIYANDO RAJAGUKGUK
221112489

UNDERGRADUATE PROGRAM OF COMPUTER SCIENCE
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026

LEMBARAN PENGESAHAN

DETEKSI PHISHING MENGGUNAKAN INFORMATION GAIN DAN RANDOM FOREST

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

ALFAJAR ALVIN PERMANA TAMBUNAN
NIM: 221112970
DWI AMANDA MONICA SIMANJUNTAK
NIM: 221113275
RIYANDO RAJAGUKGUK
NIM: 221112489

Disetujui Oleh:

Dosen Pembimbing I,	Dosen Pembimbing II,
Andri, S. Kom., M.T.I.	Kelvin, S. Kom., M. Kom.

Medan, 31 Juli 2026
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika,

Carles Juliandy, S. Kom., M. Kom

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Alfajar Alvin Permana Tambunan
NIM : 221112970

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Deteksi Phishing Menggunakan Information Gain Dan Random Forest
Tempat Penelitian : Universitas Mikroskil
Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Kota Medan
No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 1 Juli 2026
Saya yang membuat pernyataan,

Alfajar Alvin Permana Tambunan

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Riyando Rajagukguk
NIM : 221112489

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Deteksi Phishing Menggunakan Information Gain Dan Random Forest
Tempat Penelitian : Universitas Mikroskil
Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Kota Medan
No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 1 Juli 2026
Saya yang membuat pernyataan,

Riyando Rajagukguk

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Dwi Amanda Monica Simanjuntak
NIM : 221113275

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Deteksi Phishing Menggunakan Information Gain Dan Random Forest
Tempat Penelitian : Universitas Mikroskil
Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Kota Medan
No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 1 Juli 2026
Saya yang membuat pernyataan,

Dwi Amanda Monica Simanjuntak

DETEKSI PHISHING MENGGUNAKAN INFORMATION GAIN DAN RANDOM FOREST

ABSTRAK

Phishing merupakan salah satu bentuk serangan siber yang memanfaatkan website palsu untuk memperoleh informasi sensitif pengguna. Meningkatnya jumlah serangan *phishing* mendorong perlunya sistem deteksi yang memiliki akurasi tinggi serta efisiensi komputasi. Penelitian ini bertujuan menganalisis efektivitas metode Information Gain sebagai seleksi fitur pada algoritma Random Forest dengan membandingkan performa model sebelum dan sesudah seleksi fitur, serta mengimplementasikan model terbaik ke dalam browser extension. Penelitian menggunakan dua dataset, yaitu PhiUSIIL *Phishing* URL Dataset dan *Phishing* Detection Dataset (Mendeley). Tahapan penelitian meliputi preprocessing data, seleksi fitur menggunakan Information Gain, pembagian data dengan rasio 70:30, pelatihan model Random Forest, serta evaluasi menggunakan Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix. Hasil penelitian menunjukkan bahwa Information Gain mampu mengurangi jumlah fitur yang digunakan sehingga menghasilkan model yang lebih sederhana dan efisien tanpa menurunkan performa klasifikasi secara signifikan dibandingkan Random Forest tanpa seleksi fitur. Model terbaik kemudian diimplementasikan pada browser extension yang mampu mendeteksi website *phishing* secara otomatis dan memberikan peringatan kepada pengguna ketika mengakses website yang terindikasi *phishing*.

Kata Kunci: *Phishing, Information Gain, Random Forest, Seleksi Fitur, Browser Extension.*

ABSTRACT

Phishing is a form of cyberattack that uses fraudulent websites to obtain users' sensitive information. The increasing number of phishing attacks highlights the need for a detection system with high accuracy and computational efficiency. This study aims to analyze the effectiveness of the Information Gain feature selection method on the Random Forest algorithm by comparing the model performance before and after feature selection, and to implement the best-performing model into a browser extension. Two datasets were used in this study the PhiUSIIL Phishing URL Dataset and the Phishing Detection Dataset from Mendeley. The research process consisted of data preprocessing, feature selection using Information Gain, data splitting with a 70:30 ratio, Random Forest model training, and performance evaluation using Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. The results indicate that Information Gain successfully reduces the number of features, producing a simpler and more efficient model without significantly decreasing classification performance compared to the Random Forest model without feature selection. The best model was successfully implemented as a browser extension capable of automatically detecting phishing websites and providing warning notifications to users.

Keywords: *Phishing, Information Gain, Random Forest, Feature Selection, Browser Extension.*

KATA PENGANTAR

Ucapan syukur kami panjatkan kehadirat Allah SWT, Tuhan Yang Maha Esa, atas segala rahmat, hidayah, dan karunia-Nya yang melimpah, sehingga kami bertiga, ALFAJAR ALVIN PERMANA TAMBUNAN, DWI AMANDA MONICA SIMANJUNTAK, dan RIYANDO RAJAGUKGUK, dapat menyelesaikan penyusunan Skripsi ini dengan judul **" Deteksi Phishing Menggunakan Information Gain Dan Random Forest "** tepat waktu dan tanpa hambatan yang berarti. Skripsi ini disusun sebagai pemenuhan salah satu syarat akademik untuk mencapai gelar Sarjana Teknik Informatika di Universitas Mikroskil Medan. Penelitian ini memiliki urgensi tinggi karena menanggapi peningkatan kasus *phishing* yang menargetkan sektor finansial digital di Indonesia, yang berpotensi menimbulkan kerugian besar bagi pengguna lokal. Sebagai kontribusi di bidang keamanan siber, Skripsi ini berupaya mengembangkan solusi praktis dan efisien berupa extension browser yang dapat melindungi pengguna secara real-time dengan metode deteksi yang ringan.

Penyelesaian Skripsi ini tidak terlepas dari dukungan moral, materi, dan bimbingan dari banyak pihak. Oleh karena itu, dengan kerendahan hati kami menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Andri, S. Kom, M.T.I, selaku Dosen Pembimbing I.
2. Bapak Kelvin, S. Kom, M. Kom, selaku Dosen Pembimbing II.
3. Bapak Hardy, S. Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S. Kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Carles Juliandy, S. Kom., M. Kom., selaku Ketua Program Studi S-1 Teknik Informatika Fakultas Informatika Universitas Mikroskil Medan.
6. Ayahanda dan Ibunda tercinta, serta seluruh keluarga besar, atas doa, dukungan moral, dan motivasi yang tak pernah putus.
7. Seluruh rekan-rekan seperjuangan Program Studi Teknik Informatika, angkatan 2022, atas diskusi, kritik membangun, dan semangat kebersamaan selama masa perkuliahan.

Kekurangan dan kelebihan Kami menyadari bahwa Skripsi ini, meskipun telah diupayakan secara maksimal, masih memiliki kekurangan baik dari sisi substansi, metodologi, maupun tata bahasa. Namun, kami berharap kelebihan dari penelitian ini, yaitu

pengembangan prototipe extension browser yang spesifik untuk keamanan finansial dengan pendekatan heuristic yang efisien, dapat memberikan kontribusi nyata di bidang keamanan web dan menjadi referensi awal yang berguna bagi penelitian selanjutnya.

Medan, 1 Juli 2026

Penulis,

Alfajar Alvin Permana Tambunan

Dwi Amanda Monica SImanjuntak

Riyando Rajagukguk

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL