



**IMPLEMENTASI SELEKSI FITUR ALGORITMA GENETIKA PADA  
KLASIFIKASI SITUS *PHISING* DENGAN *NAÏVE BAYES***

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**MUHAMMAD YASIR HABIBIE**

**NIM 1711016310021**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**JUNI 2024**



**IMPLEMENTASI SELEKSI FITUR ALGORITMA GENETIKA PADA  
KLASIFIKASI SITUS *PHISING* DENGAN *NAÏVE BAYES***

**Skripsi**

**Untuk Memenuhi Persyaratan  
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

**Oleh**

**MUHAMMAD YASIR HABIBIE**

**NIM 1711016310021**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
JUNI 2024**

SKRIPSI

IMPLEMENTASI SELEKSI FITUR ALGORITMA GENETIKA PADA  
KLASIFIKASI SITUS *PHISING* DENGAN *NAÏVE BAYES*

Oleh :

MUHAMMAD YASIR HABIBIE

1711016310021

Telah dipertahankan di depan Dosen Penguji pada tanggal 26 Juni 2024

Susunan Dosen Penguji :

**Pembimbing I**



Muhammad Itqan Mazdadi, S.Kom., M.Kom

NIP. 199006122019031013

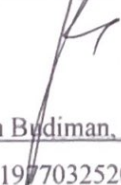
**Dosen Penguji I**



Triando H. Saragih, S.Kom., M.Kom

NIP. 19930824201931012

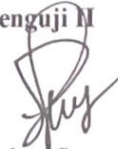
**Pembimbing II**



Irwan Budiman, S.T., M.Kom

NIP. 197703252008121001

**Dosen Penguji II**



Setyo Wahyu Saputro, S.Kom, M.Kom

NIP. 198808072023211027

Banjarbaru, 27 Juni 2024

Ketua Program Studi Ilmu Komputer



Irwan Budiman, S.T., M.Kom

NIP. 197703252008121001

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 26 Juni 2024

Yang Menyatakan,



Muhammad Yasir Habibie  
NIM. 1711016310021

## ABSTRAK

### **IMPLEMENTASI SELEKSI FITUR ALGORITMA GENETIKA PADA KLASIFIKASI SITUS PHISING DENGAN NAÏVE BAYES**

(Oleh: Muhammad Yasir Habibie; Pembimbing: Muhammad Itqan Mazdadi, S.Kom., M.Kom dan Irwan Budiman, S.T., M.Kom; 2024; 42 halaman)

*Phising merupakan penipuan yang menggunakan manipulasi psikologis untuk mencuri informasi sensitif seperti kata sandi atau informasi pribadi lainnya yang merugikan bagi pengguna internet. Penggunaan teknik-teknik kecerdasan buatan seperti machine learning dinilai sangat efektif dalam mendeteksi situs phishing dengan tingkat akurasi yang tinggi. Naïve Bayes merupakan salah satu algoritma klasifikasi yang populer dalam machine learning dengan kelebihan utama yaitu kesederhanaannya yang memungkinkan untuk dilatih dengan cepat bahkan dengan dataset yang besar, sehingga cocok untuk diaplikasikan pada sumber daya yang terbatas. Meski demikian, Naïve Bayes memiliki kelemahan yaitu sangat sensitif dalam pemilihan fitur, oleh karena itu algoritma genetika digunakan untuk menutupi kelemahan dari Naïve Bayes dengan mencari kombinasi fitur terbaik yang dapat meningkatkan akurasi prediksi dari Naïve Bayes. Hasil penelitian menunjukkan bahwa seleksi fitur menggunakan algoritma genetika pada klasifikasi situs phishing menghasilkan akurasi sebesar 93,08%. Seleksi fitur dengan algoritma genetika dapat memberikan peningkatan kinerja yang signifikan pada model Naïve Bayes dengan peningkatan akurasi sebesar 31,34% dibandingkan dengan penggunaan seluruh fitur tanpa seleksi.*

**Kata kunci:** *Phising, Naïve Bayes, Seleksi Fitur, Algoritma Genetika.*

## **ABSTRACT**

### **IMPLEMENTATION OF GENETIC ALGORITHM FEATURE SELECTION IN PHISHING WEBSITE CLASSIFICATION USING NAÏVE BAYES**

*(By: Muhammad Yasir Habibie; Supervisor: Muhammad Itqan Mazdadi, S.Kom., M.Kom and Irwan Budiman, S.T., M.Kom; 2024; 42 pages)*

*Phishing is a fraud that uses psychological manipulation to steal sensitive information such as passwords or other personal information that is detrimental to internet users. The use of artificial intelligence techniques such as machine learning is considered very effective in detecting phishing sites with a high level of accuracy. Naïve Bayes is a popular classification algorithm in machine learning with the main advantage being its simplicity which allows it to be trained quickly even with large datasets, making it suitable for application on limited resources. However, Naïve Bayes has a weakness, namely that it is very sensitive in feature selection, therefore genetic algorithms are used to cover the weaknesses of Naïve Bayes by looking for the best combination of features that can increase the prediction accuracy of Naïve Bayes. The research results show that feature selection using a genetic algorithm for classifying phishing sites produces an accuracy of 93.08%. Feature selection with a genetic algorithm can provide a significant performance increase in the Naïve Bayes model with an increase in accuracy of 31.34% compared to using all features without selection.*

**Keywords:** *Phishing, Naïve Bayes, Feature Selection, Genetic Algorithm.*

## PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Seleksi Fitur Algoritma Genetika Pada Klasifikasi Situs *Phising* dengan *Naïve Bayes*” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan shalawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga *yaumul qiyamah*.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

1. Keluarga terutama kepada Orang tua yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian skripsi ini.
2. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran penyelesaian skripsi ini.
3. Bapak Irwan Budiman, S.T., M.Kom selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Bapak Irwan Budiman, S.T., M.Kom selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
6. Teman-teman keluarga Ilmu Komputer angkatan 2017, terima kasih untuk canda, tawa, perjuangan yang sudah dilewati bersama, untuk semua kenangan manis yang telah terukir selama ini. Senang bisa menjadi salah satu bagian dari kehidupan kalian.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini.

Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Juni 2024

A handwritten signature in blue ink, appearing to read 'M. YASIR HABIBIE', with a stylized flourish at the end.

Muhammad Yasir Habibie

## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN .....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan.....	3
1.5 Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Kajian Terdahulu .....	4
2.2 Landasan Teori .....	8
2.2.1 <i>Phising</i> .....	8
2.2.2 <i>Phishing Detection</i> .....	9
2.2.3 Seleksi Fitur .....	15
2.2.4 Algoritma Genetika.....	16
2.2.5 <i>Fitness Function</i> .....	20

2.2.6	<i>Naïve Bayes</i> .....	20
2.2.7	<i>Confusion matrix</i> .....	21
BAB III METODE PENELITIAN.....		23
3.1	Bahan Penelitian .....	23
3.2	Alat Penelitian .....	23
3.3	Variabel Penelitian.....	23
3.4	Prosedur Penelitian .....	23
BAB IV HASIL DAN PEMBAHASAN .....		26
4.1	Hasil.....	26
4.1.1	Pengumpulan <i>Dataset</i> .....	26
4.1.2	<i>Modeling</i> .....	29
4.1.3	Evaluasi Hasil Kinerja Klasifikasi .....	45
4.2	Pembahasan .....	47
BAB V PENUTUP.....		51
5.1	Kesimpulan.....	51
5.2	Saran .....	51
DAFTAR PUSTAKA .....		52
LAMPIRAN.....		54

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
Tabel 1 Keaslian Penelitian.....	5
Tabel 2 Perancangan Penelitian .....	8
Tabel 3 Contoh Inisialiasai populasi .....	17
Tabel 4 Contoh <i>crossover</i> pada algoritma genetika .....	19
Tabel 5 <i>Confusion matrix Binary Classification</i> .....	22
Tabel 6 Fitur pada <i>Phising Website Data Set</i> .....	26
Tabel 7 Contoh <i>dataset Phising Website Data Set</i> .....	28
Tabel 8 Parameter Algoritma Genetika.....	32
Tabel 9. Inisialisasi Populasi Algoritma Genetika .....	33
Tabel 10 Nilai <i>Fitness</i> Kromosom .....	34
Tabel 11 Ilustrasi Seleksi Turnamen pada Iterasi Pertama .....	36
Tabel 12 Populasi Baru Algoritma Genetika .....	37
Tabel 13 Ilustrasi <i>Crossover</i> Algoritma Genetika pada Iterasi Pertama.....	38
Tabel 14 Ilustrasi <i>Crossover</i> Algoritma Genetika pada Individu Ke-20 .....	39
Tabel 15 Populasi setelah <i>crossover</i> .....	39
Tabel 16 Ilustrasi Pemilihan Gen yang Dimutasi pada Iterasi Pertama.....	40
Tabel 17 Ilustrasi Mutasi Gen pada Individu Ke-2 .....	41
Tabel 18 Populasi Setelah Proses Mutasi .....	41
Tabel 19 Kromosom terbaik yang dihasilkan dari setiap iterasi .....	42
Tabel 20 Perhitungan akurasi klasifikasi <i>Naïve Bayes</i> tanpa <i>seleksi fitur</i> .....	45
Tabel 21 Perhitungan akurasi klasifikasi <i>Naïve Bayes</i> dengan Seleksi Fitur.....	46
Tabel 22 Fitur yang Dihasilkan dari Seleksi Fitur Algoritma Genetika .....	47

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
Gambar 1 Anatomi <i>URL</i> .....	10
Gambar 2 Model hierarki prediksi situs <i>phising</i> .....	11
Gambar 3 <i>Flowchart Naïve Bayes</i> .....	21
Gambar 4 Alur Penelitian.....	24
Gambar 5 Distribusi kelas pada <i>Phising Website Data Set</i> .....	28
Gambar 6 Perbandingan persentase kelas <i>phising</i> dengan kelas non <i>phising</i> .....	29
Gambar 7 Pembagian <i>Dataset</i> .....	29
Gambar 8 Alur Seleksi Fitur dengan Algoritma Genetika .....	31
Gambar 9 Grafik Seleksi Fitur Algoritma Genetika .....	44
Gambar 10 <i>Confusion matrix</i> klasifikasi <i>Naïve Bayes</i> tanpa seleksi fitur .....	45
Gambar 11 <i>Confusion matrix</i> klasifikasi <i>Naïve Bayes</i> dengan seleksi fitur .....	46
Gambar 12 Grafik Perbandingan Kinerja Model .....	49

## DAFTAR LAMPIRAN

### **Lampiran**

Lampiran 1 *Source Code* Import data dan pembagian *dataset*

Lampiran 2 *Source Code* Klasifikasi *Naïve Bayes* tanpa Seleksi Fitur

Lampiran 3 *Source Code* Seleksi Fitur Algoritma Genetika

Lampiran 4 *Source Code* Klasifikasi *Naïve Bayes* dengan Seleksi Fitur Algoritma Genetika