

**OPTIMISASI MODEL MACHINE LEARNING UNTUK
DETEKSI SPAM EMAIL: PERBANDINGAN KINERJA
ALGORITMA NAÏVE BAYES DAN ALGORITMA C45**

SKRIPSI SARJANA REKAYASA TEKNOLOGI INFORMATIKA

Oleh:

Muhammad Luthfi Prasetyo

207064516019



PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA

UNIVERSITAS NASIONAL

2023

**OPTIMISASI MODEL MACHINE LEARNING UNTUK
DETEKSI SPAM EMAIL: PERBANDINGAN KINERJA
ALGORITMA NAÏVE BAYES DAN ALGORITMA C45**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
Teknologi Informasi dan Fakultas Teknologi Komunikasi dan Informatika

Oleh:

Muhammad Luthfi Prasetyo

207064516019



PROGRAM STUDI INFORMATIKA

**FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA**

UNIVERSITAS NASIONAL

2020

HALAMAN PENGESAHAN

TUGAS AKHIR

OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM
EMAIL: PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN
ALGORITMA C45



Muhammad Luthfi Prasetyo

207064516019

Dosen Pembimbing 1

A handwritten signature in black ink, consisting of stylized, overlapping letters.

(Dr. Agung Triayudi, S.Kom. ,M.Kom)

Dosen Pembimbing 2

A handwritten signature in black ink, featuring a prominent 'G' and 'S'.

(Gatot Soepriyono, S.Si., M.S.M.)

HALAMAN PERNYATAAN ORISINALITAS

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bilamana di kemudian hari ditemukan bahwa karya tulis ini menyalahi peraturan yang ada berkaitan etika dan kaidah penulisan karya ilmiah yang berlaku, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Yang menyatakan,

Nama : Muhammad Luthfi Prasetyo

NIM : 207064516019

Tanda Tangan 

Tanggal : 23 Agustus 2024

Mengetahui

Pembimbing 1 : Dr. Agung Triayudi, S.Kom., M.Kom. ()

Pembimbing 2 : Gatot Soepriyono, S.Si., M.S.M. ()



HALAMAN PENGESAHAN

TUGAS SARJANA

HALAMAN PENGESAHAN

TUGAS SARJANA

OPTIMISASI MODEL MACHINE LEARNING UNTUK
DETEKSI SPAM EMAIL: PERBANDINGAN KINERJA
ALGORITMA NAÏVE BAYES DAN ALGORITMA C45

Oleh:

Muhammad Luthfi Prasetyo

207064516019

Program Studi Informatika

Fakultas Teknologi Komunikasi dan Informatika

Disetujui pada tanggal: 23 Agustus 2024

Pembimbing 1



Dr. Agung Triayudi, S.Kom., M.Kom.

NIDN. 0419068604

Penguji 1



Dr. Andrianingsih, S.Kom., MMSI

NIDN. 0303097902

Pembimbing II



Gatot Soepriyono, S.Sr., M.S.M.

NIDN. 0317047803

Penguji 2



Agus Iskandar, S.kom., M.Kom.

NIDN. 0310087503

UNIVERSITAS NASIONAL

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM EMAIL PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN ALGORITMA C

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasionl sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasionl maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 23 Agustus 2023



Muhammad Luthfi Prasetyo

2070645160

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul:

OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM EMAIL: PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN ALGORITMA C45

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Akhir Semester Genap 2023 – 2024 pada tanggal 21 Agustus 2024



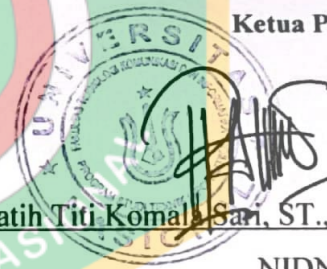
Dosen Pembimbing 1

A handwritten signature in black ink, appearing to be 'A. Triayudi', written over a light blue background.

Dr. Agung Triayudi, S.Kom., M.Kom

NIDN 0327047903

Ketua Program Studi



Ratih Titi Komala Sari, ST., MM., MMSI

NIDN 0301038302

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Muhammad Luthfi Prasetyo
NPM : 207064516019
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Informatika
Tanggal Sidang : 21 Agustus 2024



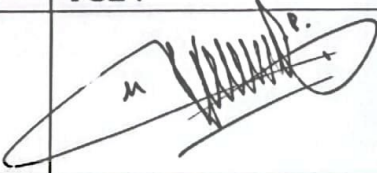
JUDUL DALAM BAHASA INDONESIA :

**OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM
EMAIL: PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN
ALGORITMA C45**

JUDUL DALAM BAHASA INGGRIS :

**OPTIMIZATION OF MACHINE LEARNING MODELS FOR EMAIL SPAM
DETECTION: PERFORMANCE COMPARISON OF NAIVE BAYES ALGORITHM
AND C45 ALGORITHM**

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL :	TGL :	TGL :
		

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Muhammad Luthfi Prasetyo
NPM : 207064516019
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Informatika
Tanggal Sidang : 21 Agustus 2024



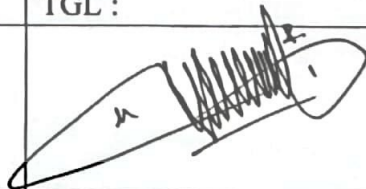
JUDUL DALAM BAHASA INDONESIA :

**OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM
EMAIL: PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN
ALGORITMA C45**

JUDUL DALAM BAHASA INGGRIS :

**OPTIMIZATION OF MACHINE LEARNING MODELS FOR EMAIL SPAM
DETECTION: PERFORMANCE COMPARISON OF NAIVE BAYES ALGORITHM
AND C45 ALGORITHM**

TANDA TANGAN DAN TANGGAL

Pembimbing 2	Ka. Prodi	Mahasiswa
TGL :	TGL :	TGL :
		

KATA PENGANTAR`

Puji Syukur penulis panjatkan atas kehadiran Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul **“OPTIMISASI MODEL MACHINE LEARNING UNTUK DETEKSI SPAM EMAIL: PERBANDINGAN KINERJA ALGORITMA NAÏVE BAYES DAN ALGORITMA C45”**. Adapun tujuan dari penyusunan skripsi ini ialah untuk memenuhi salah satu syarat memperoleh gelar Sarjana Komputer di Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional Jakarta.

Dalam penyusunan skripsi ini penulis banyak menerima bantuan, bimbingan serta dukungan dari berbagai pihak, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Maka dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dalam penyusunan skripsi ini, terutama kepada yang terhormat :

1. Bapak Dr. Agung Triayudi, S.Kom.,M.Kom Selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional
2. Ibu Ir. Endah Tri Esti Handayani, MMSI.. Selaku Wakil Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional
3. Ibu Ratih Titi Komalasari, ST, MM, MMSI Selaku Ketua Program Studi Informatika Universitas Nasional
4. Ibu Rima Tamara Aldisa, S.Kom., M.Kom.Selaku Sekretaris Program Studi Informatika Universitas Nasional
5. Para Dosen dan Seluruh Staff akademik Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional yang telah memberikan bekal ilmu yang bermanfaat.
6. Kedua orangtua tercinta dan saya sayangi yang senantiasa mencurahkan segenap kasih sayang yang tiada henti-hentinya, doa, motivasi, nasehat, serta kesabaran yang begitu besar.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak

kekurangan baik dalam segi materi, isi maupun teknik penyajiannya, mengingat kemampuan yang penulis miliki masih terbatas. Sehingga kritik dan saran yang bersifat membangun dari semua pihak senantiasa dengan senang hati penulis harapkan demi kesempurnaan penyusunan skripsi kemudian.

Jakarta, Agustus 2024

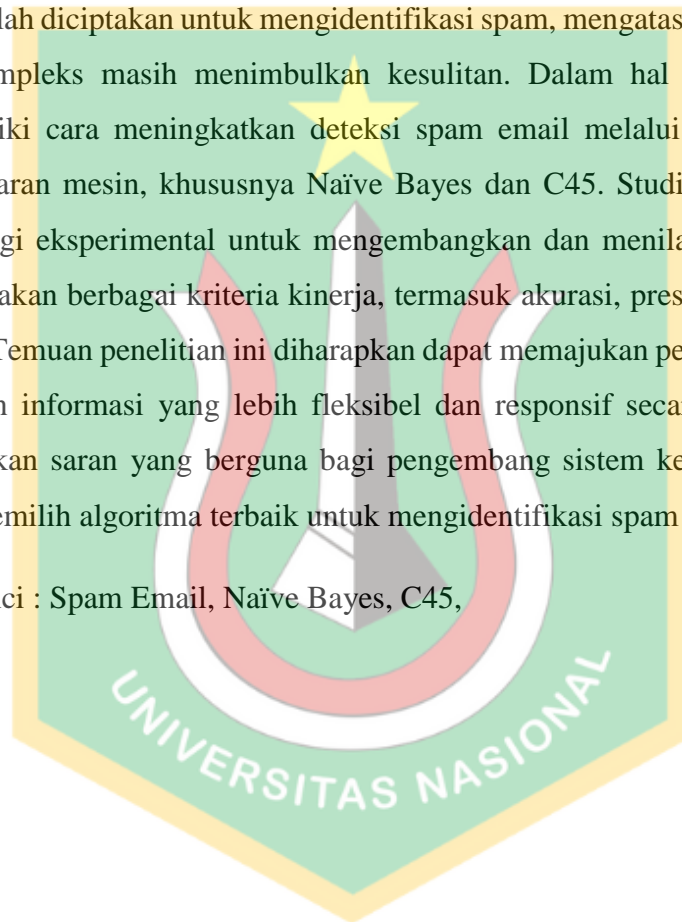


Penulis

ABSTRAK

Email telah berkembang menjadi alat komunikasi yang vital di era digital yang berkembang pesat, namun juga rentan terhadap bahaya spam email. Email spam, yang sering kali berisi konten yang tidak diminta seperti penawaran palsu dan tautan berbahaya, sangat membahayakan keamanan informasi dan produktivitas. Meskipun sejumlah strategi, seperti pendekatan heuristik dan filter berbasis aturan, telah diciptakan untuk mengidentifikasi spam, mengatasi teknik spam yang lebih kompleks masih menimbulkan kesulitan. Dalam hal ini, penelitian ini menyelidiki cara meningkatkan deteksi spam email melalui penerapan teknik pembelajaran mesin, khususnya Naïve Bayes dan C45. Studi ini menggunakan metodologi eksperimental untuk mengembangkan dan menilai kedua algoritma menggunakan berbagai kriteria kinerja, termasuk akurasi, presisi, perolehan, dan skor F1. Temuan penelitian ini diharapkan dapat memajukan penciptaan teknologi keamanan informasi yang lebih fleksibel dan responsif secara signifikan serta memberikan saran yang berguna bagi pengembang sistem keamanan informasi dalam memilih algoritma terbaik untuk mengidentifikasi spam email.

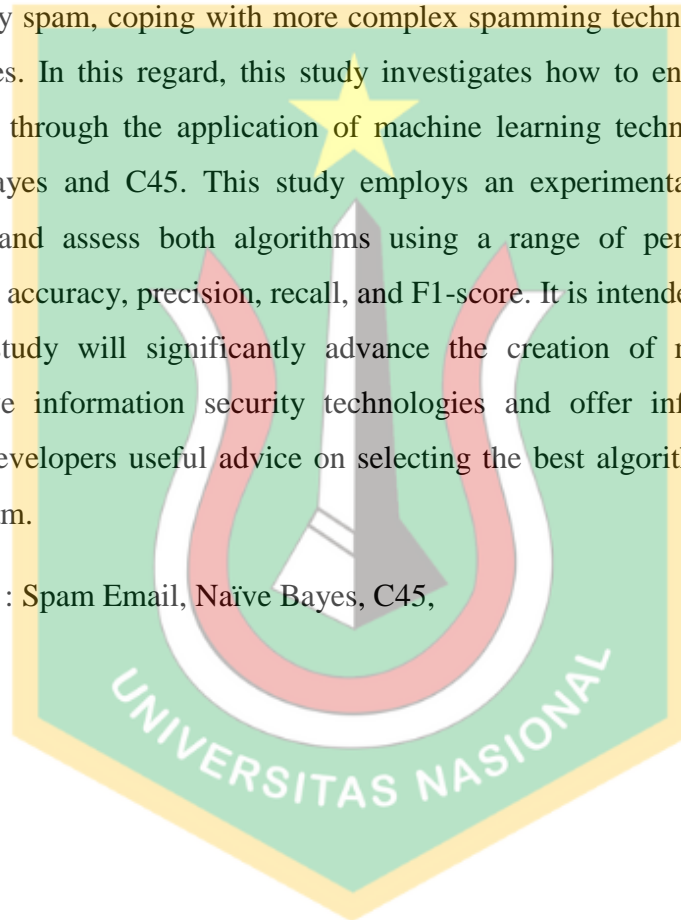
Kata Kunci : Spam Email, Naïve Bayes, C45,



ABSTRACT

Email has grown to be a vital tool for communication in the rapidly expanding digital age, but it is also susceptible to the danger of email spam. Spam emails, which frequently contain unsolicited content like bogus offers and harmful links, seriously jeopardize information security and productivity. Even though a number of strategies, such as heuristic approaches and rule-based filters, have been created to identify spam, coping with more complex spamming techniques still presents difficulties. In this regard, this study investigates how to enhance email spam detection through the application of machine learning techniques, particularly Naïve Bayes and C45. This study employs an experimental methodology to develop and assess both algorithms using a range of performance criteria, including accuracy, precision, recall, and F1-score. It is intended that the findings of this study will significantly advance the creation of more flexible and responsive information security technologies and offer information security system developers useful advice on selecting the best algorithm for identifying email spam.

Keyword : Spam Email, Naïve Bayes, C45,



DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	i
HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR`	i
ABSTRAK	iii
ABSTRACT.....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	vii
DAFTAR TABEL.....	viii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Batasan Masalah.....	4
1.5 Kontribusi Penelitian.....	5
BAB II.....	6
TINJAUAN PUSTAKA.....	6
2.1 Machine Learning	6
2.1.1 Algoritma Naïve bayes.....	6
2.1.2 Algoritma C45.....	7
2.1.3 Draw.io.....	8
2.2 Studi Literatur	9
BAB III.....	15
METODOLOGI PENELITIAN.....	15
3.1 Waktu Penelitian.....	15
3.2 Penentuan Subjek Penelitian	15
3.3 Sumber Data.....	16
3.4 Tahapan Penelitian	17
3.5 Prototype	20
BAB IV	23

HASIL DAN PEMBAHASAN	23
4.1 Pengambilan Data	23
4.2 Perhitungan Manual	23
4.2.1 Algoritma Naïve Bayes	23
4.2.2 Algoritma C45.....	25
4.3 Pembahasan.....	29
4.3.1 Implementasi Metode Naïve Bayes	29
4.3.2 Implementasi Metode C45	29
4.4 Hasil Perbandingan	29
4.5 Implementasi Aplikasi.....	31
4.5.1 Halaman Login.....	31
4.5.2 Halaman Dashboard	32
4.5.3 Halaman Setelah Deteksi	32
BAB V.....	34
KESIMPULAN DAN SARAN.....	34
5.1 Kesimpulan	34
5.2 Saran.....	34
DAFTAR PUSTAKA	35
LAMPIRAN.....	37



DAFTAR GAMBAR

Gambar 3. 1 Flowchart Tahapan Penelitian	17
Gambar 3. 2 Flowchart Aplikasi	20
Gambar 3. 3 Mockup Login	21
Gambar 3. 4 Dashboard Ham	21
Gambar 3. 5 Dashboard Spam.....	22
Gambar 4. 1 Dataset	23
Gambar 4. 2 Dataset Naive Bayes.....	24
Gambar 4. 3 Dataset C45	25
Gambar 4. 4 Total Seluruh Data.....	26
Gambar 4. 5 Total entropy Seluruh Data.....	26
Gambar 4. 6 Nilai Setiap Atribut.....	26
Gambar 4. 7 Total Entropy Setiap Atribut.....	27
Gambar 4. 8 Total Info Gain	27
Gambar 4. 9 Total Split Info.....	28
Gambar 4. 10 Total Gain Info	28
Gambar 4. 11 Accuracy Naive Bayes.....	29
Gambar 4. 12 Accuracy C45	29
Gambar 4. 13 Grafik Total keseluruhan	30
Gambar 4. 14 Grafik Hasil Perbandingan	30
Gambar 4. 15 Tampilan Login	31
Gambar 4. 16 Halaman Dashboard	32
Gambar 4. 17 Deteksi Ham.....	32
Gambar 4. 18 Deteksi Spam.....	33

DAFTAR TABEL

Tabel 3. 1 Waktu Penelitian.....15

