

**PENGEMBANGAN MODUL U-TAPIS UNTUK DETEKSI KASUS
HARASSMENT PENGHINAAN DAN PENCEMARAN NAMA BAIK
DENGAN NAIVE BAYES CLASSIFIER**



LAPORAN MBKM PENELITIAN

**GERALD IMANUEL WIJAYA
00000060106**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**PENGEMBANGAN MODUL U-TAPIS UNTUK DETEKSI KASUS
HARASSMENT PENGHINAAN DAN PENCEMARAN NAMA BAIK
DENGAN NAIVE BAYES CLASSIFIER**



LAPORAN MBKM PENELITIAN

**GERALD IMANUEL WIJAYA
0000060106**

**UNIVERSITAS
MULTIMEDIA
NUSANTARA**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Gerald Imanuel Wijaya

NIM : 00000060106

Program Studi : Informatika

Menyatakan dengan sesungguhnya bahwa Tesis/Skripsi/Tugas Akhir/Laporan Magang/MBKM saya yang berjudul:

Pengembangan Modul U-Tapis untuk Deteksi Kasus *Harassment* Penghinaan dan Pencemaran Nama Baik dengan *Naive Bayes classifier*

merupakan hasil karya saya sendiri, bukan merupakan hasil plagiat, dan tidak pula dituliskan oleh orang lain; Semua sumber, baik yang dikutip maupun dirujuk, telah saya cantumkan dan nyatakan dengan benar pada bagian Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan karya ilmiah, saya bersedia menerima konsekuensi untuk dinyatakan TIDAK LULUS. Saya juga bersedia menanggung segala konsekuensi hukum yang berkaitan dengan tindak plagiarisme ini sebagai kesalahan saya pribadi dan bukan tanggung jawab Universitas Multimedia Nusantara.

Tangerang, 3 Januari 2025



(Gerald Imanuel Wijaya)

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Gerald Imanuel Wijaya
NIM : 00000060106
Program Studi : Informatika
Jenjang : S1
Jenis Karya : Laporan MBKM Penelitian

Menyatakan dengan sesungguhnya bahwa:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 3 Januari 2025

Yang menyatakan

Gerald Imanuel Wijaya

** Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

”I can do all things through Christ who strengthens me.”

Philippians 4:13

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Laporan MBKM Penelitian ini dengan judul: Pengembangan Modul U-Tapis untuk Deteksi Kasus Harassment Penghinaan dan Pencemaran Nama Baik dengan Naive Bayes classifier dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan laporan magang ini, sangatlah sulit bagi saya untuk menyelesaikan laporan magang ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

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 Assoc. Prof. Arya Wicaksana, S.Kom., M.Eng.Sc., OCA., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Wirawan Istiono, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya penelitian ini.
5. Marlinda Vasty Overbeek, S.Kom., M.Kom. yang sudah menuntun serta membimbing saya selama penelitian ini, sehingga penelitian dapat selesai dengan baik dan tepat waktu.
6. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan penelitian ini.
7. Stevani selaku pacar saya yang terus memberikan semangat serta dukungan secara moral, sehingga penelitian ini dapat diselesaikan dengan baik dan tepat waktu.

Semoga laporan penelitian ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 Januari 2025

Gerald Imanuel Wijaya

**PENGEMBANGAN MODUL U-TAPIS UNTUK DETEKSI KASUS
HARASSMENT PENGHINAAN DAN PENCEMARAN NAMA BAIK
DENGAN NAIVE BAYES CLASSIFIER**

Gerald Imanuel Wijaya

ABSTRAK

Kasus *harassment* berbasis teks, khususnya penghinaan dan pencemaran nama baik, semakin meningkat di era digital dan membutuhkan solusi yang cepat serta akurat. Penelitian ini mengembangkan modul U-Tapis menggunakan algoritma *Naive Bayes Classifier* (NBC) untuk mendeteksi *harassment* secara otomatis pada teks berbahasa Indonesia. Dengan dataset pelatihan sebanyak 18.000 data dan evaluasi pada 320 data uji GPT dan 1177 data uji Internet dan Twitter, model mencapai akurasi hingga 86%, dengan performa terbaik pada kategori Netral dan Pencemaran Nama Baik, namun tantangan ditemukan pada kategori Hinaan. Temuan ini menunjukkan perlunya pengayaan dataset, serta penggunaan algoritma yang lebih kompleks seperti *deep learning*. Modul ini memiliki potensi untuk diintegrasikan dengan sistem pelaporan daring, memberikan kontribusi pada otomatisasi berbasis *Natural Language Processing* (NLP) yang lebih adaptif dan efisien.

Kata kunci: *Harassment, Naive Bayes Classifier, Natural Language Processing, Pencemaran Nama Baik, Penghinaan*

***THE DEVELOPMENT OF THE U-TAPIS MODULE FOR DETECTING
HARASSMENT AND DEFAMATION CASES OF INSULTS USING THE
NAIVE BAYES CLASSIFIER***

Gerald Imanuel Wijaya

ABSTRACT

Text-based harassment cases, particularly insults and defamation, are increasingly prevalent in the digital era, requiring fast and accurate solutions. This study developed the U-Tapis module using the Naive Bayes Classifier (NBC) algorithm to automatically detect harassment in Indonesian texts. With a training dataset of 18,000 entries and evaluation on 320 GPT test data and 1177 Internet and Twitter test data, the model achieved an accuracy of up to 86%, performing best in the Neutral and Defamation categories, although challenges remain in the Insult category. These findings highlight the need for dataset enrichment and the adoption of more complex algorithms, such as deep learning. This module has the potential to be integrated with online reporting systems, contributing to more adaptive and efficient Natural Language Processing (NLP)-based automation.

Keywords: *Defamation, Harassment, Insult, Naive Bayes Classifier, Natural Language Processing*

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iii
HALAMAN PERSEMBAHAN/MOTO	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Urgensi Penelitian	4
1.5 Luaran Penelitian	4
1.6 Manfaat Penelitian	5
BAB 2 TINJAUAN PUSTAKA	6
2.1 <i>Harassment</i>	6
2.2 Penghinaan	6
2.3 Pencemaran Nama Baik	6
2.4 Algoritma dan Metode	7
2.4.1 <i>Natural Language Processing</i>	7
2.4.2 <i>Text Preprocessing</i>	7
2.4.3 <i>Supervised Learning</i>	7
2.5 <i>Naive Bayes Classifier</i>	8
2.5.1 <i>Multinomial Naive Bayes</i>	9
BAB 3 METODE PENELITIAN	11
3.1 Identifikasi Masalah	12
3.2 Studi Literatur	12
3.3 Pengumpulan Data	12
3.4 Pembuatan Model	13
3.4.1 <i>Text Cleaning and Preprocessing</i>	14
3.4.2 <i>Feature Extraction</i>	15
3.4.3 Pembagian Dataset	16
3.4.4 Pelatihan Model Naive Bayes	16
3.4.5 Evaluasi Model	17
3.4.6 <i>Hyperparameter Tuning</i> menggunakan GridSearchCV	18
3.4.7 Pengujian Model dengan Data Testing	18
3.5 Evaluasi Hasil Uji Coba	18
3.6 Implementasi Model Pembelajaran Mesin ke Aplikasi <i>Website</i>	18
BAB 4 HASIL DAN PEMBAHASAN	20
4.1 Hasil Perancangan	20
4.1.1 Hasil Implementasi Backend	20
4.2 Pembahasan Hasil Rancangan	21
4.2.1 Implementasi Sistem Deteksi Pencemaran Nama Baik dan Hinaan	21

4.2.2	Pengujian Sistem	25
4.2.3	Ringkasan Hasil Akurasi pada Setiap Dataset yang Diuji . .	34
BAB 5	SIMPULAN SARAN	36
5.1	Simpulan	36
5.2	Saran	36
DAFTAR PUSTAKA	38

DAFTAR GAMBAR

Gambar 3.1	<i>Research Methodology Flowchart</i>	11
Gambar 3.2	<i>Model Training Flowchart</i>	14
Gambar 3.3	<i>Features Extraction Flowchart</i>	15
Gambar 3.4	Pelatihan Model <i>Naive Bayes</i>	16
Gambar 4.1	Pengujian <i>Backend Request</i>	20
Gambar 4.2	<i>Import Dataset</i>	22
Gambar 4.3	<i>Slangword Mapping</i>	23
Gambar 4.4	<i>Text Preprocessing</i>	23
Gambar 4.5	Vektorisasi	24
Gambar 4.6	Pembangunan Model	24

DAFTAR TABEL

Tabel 4.1	<i>Confusion Matrix</i> untuk Data Training	26
Tabel 4.2	Laporan Klasifikasi untuk Data Training	26
Tabel 4.3	<i>Confusion Matrix</i> untuk Data Testing	26
Tabel 4.4	Laporan Klasifikasi untuk Data Testing	26
Tabel 4.5	<i>Confusioni Matrix</i> untuk Algoritma Naive Bayes (20 Sampel)	27
Tabel 4.6	Laporan Klasifikasi untuk Algoritma Naive Bayes (20 Sampel)	27
Tabel 4.7	<i>Confusioni Matrix</i> untuk Algoritma Naive Bayes (50 Sampel)	29
Tabel 4.8	Laporan Klasifikasi untuk Algoritma Naive Bayes (50 Sampel)	29
Tabel 4.9	<i>Confusioni Matrix</i> untuk Algoritma Naive Bayes (100 Sampel)	30
Tabel 4.10	Laporan Klasifikasi untuk Algoritma Naive Bayes (100 Sampel)	30
Tabel 4.11	<i>Confusion Matrix</i> untuk Algoritma Naive Bayes (150 Sampel)	32
Tabel 4.12	Laporan Klasifikasi untuk Algoritma Naive Bayes (150 Sampel)	32
Tabel 4.13	<i>Confusion Matrix</i> untuk Algoritma Naive Bayes (1177 Sampel)	33
Tabel 4.14	Laporan Klasifikasi untuk Algoritma Naive Bayes (1177 Sampel)	33
Tabel 4.15	Ringkasan Hasil Akurasi Algoritma Naive Bayes pada Setiap Dataset	35

DAFTAR LAMPIRAN

Lampiran 1	MBKM-01 Cover Letter MBKM Internship Track 1	42
Lampiran 2	MBKM-02 MBKM Internship Track 1 Card	43
Lampiran 3	MBKM-03 Daily Task - Internship Track 1	44
Lampiran 4	MBKM-04 Verification Form of Internship Report MBKM Internship Track 1	53
Lampiran 5	Form Bimbingan	54
Lampiran 6	Hasil Pemeriksaan Turnitin	55
Lampiran 7	MoU Kerjasama Penelitian	61
Lampiran 8	Draft Paper	64