

**U-TAPIS: DETEKSI DAN KOREKSI KESALAHAN
PENULISAN PELULUHAN KATA MENGGUNAKAN
ALGORITMA DAMERAU-LEVENSHTTEIN DISTANCE DAN
ROBERTA**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

**PRUDENCE TENDY
00000060765**

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

**U-TAPIS: DETEKSI DAN KOREKSI KESALAHAN
PENULISAN PELULUHAN KATA MENGGUNAKAN
ALGORITMA DAMERAU-LEVENSHTEIN DISTANCE DAN**

ROBERTA



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**PRUDENCE TENDY
00000060765**

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

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Prudence Tendy
Nomor Induk Mahasiswa : 00000060765
Program Studi : Informatika

Skripsi dengan judul:

U-Tapis: Deteksi dan Koreksi Kesalahan Penulisan Peluluhan Kata Menggunakan Algoritma Damerau-Levenshtein Distance dan RoBERTa

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 04 Maret 2025



(Prudence Tendy)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

**U-TAPIS: DETEKSI DAN KOREKSI KESALAHAN PENULISAN
PELULUHAN KATA MENGGUNAKAN ALGORITMA
DAMERAU-LEVENSHTEIN DISTANCE DAN ROBERTA**

oleh

Nama : Prudence Tendency
NIM : 00000060765
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Rabu, 12 Maret 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Pembimbing



(Marlinda Vasty Overbeek, S.Kom.,
M.Kom.)

NIDN: 0818038501

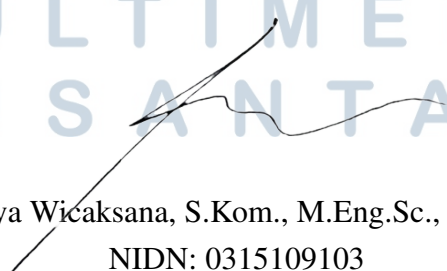
Ketua Sidang



(Dr. Adhi Kusnadi, S.T., M.Si.)

NIDN: 0303037304

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

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

Yang bertanda tangan di bawah ini:

Nama : Prudence Tendency
NIM : 00000060765
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : U-Tapis: Deteksi dan Koreksi
Kesalahan Penulisan Peluluhan Kata
Menggunakan Algoritma Damerau-
Levenshtein Distance dan RoBERTa

Menyatakan dengan sesungguhnya bahwa saya bersedia:

- ☐ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☒ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 04 Maret 2025

Yang menyatakan



Prudence Tendency

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

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: U-Tapis: Deteksi dan Koreksi Kesalahan Penulisan Peluluhan Kata Menggunakan Algoritma Damerau-Levenshtein Distance dan RoBERTa 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:

Mengucapkan terima kasih

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 Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 04 Maret 2025



Prudence Tendy

**U-TAPIS: DETEKSI DAN KOREKSI KESALAHAN PENULISAN
PELULUHAN KATA MENGGUNAKAN ALGORITMA
DAMERAU-LEVENSHTEIN DISTANCE DAN ROBERTA**

Prudence Tendency

ABSTRAK

Di era digital saat ini, permintaan akan penyampaian berita yang cepat meningkatkan risiko kesalahan linguistik, termasuk kesalahan dalam penggunaan peluluhan kata. Penelitian ini memperkenalkan aplikasi U-Tapis, sebuah platform yang dirancang untuk mendeteksi dan memperbaiki kesalahan tersebut menggunakan algoritma Damerau-Levenshtein Distance dan model RoBERTa. Sistem ini mencapai akurasi rekomendasi rata-rata sebesar 92.84%, dengan kinerja berkisar antara 91.30% hingga 95.45% pada 3000 artikel berita. Meskipun efektif, sistem ini memiliki keterbatasan, seperti sifat statis dari datasetnya yang tidak diperbarui secara dinamis dengan entri baru dalam Kamus Besar Bahasa Indonesia, serta kecenderungannya untuk menandai semua kata dengan prefiks “me-” dan “pe-” tanpa mempertimbangkan konteks. Tantangan ini menyoroti peluang untuk pengembangan lebih lanjut guna meningkatkan adaptabilitas dan ketepatan platform.

Kata kunci: Damerau-Levenshtein Distance, Peluluhan Kata, RoBERTa, U-Tapis



***U-TAPIS: A HYBRID APPROACH TO MELTING WORD ERROR
DETECTION AND CORRECTION WITH DAMERAU-LEVENSHTEIN
DISTANCE AND ROBERTA***

Prudence Tendy

ABSTRACT

In the current digital era, the demand for rapid news delivery increases the risk of linguistic errors, including inaccuracies in the usage of melting words. This study introduces the U-Tapis application, a platform designed to detect and correct such errors using the Damerau-Levenshtein Distance algorithm and the RoBERTa model. The system achieved an average recommendation accuracy of 92.84%, with performance ranging from 91.30% to 95.45% across 3000 news articles. Despite its effectiveness, the system faces limitations, such as the static nature of its dataset, which does not update dynamically with new entries in the Indonesian Language Dictionary, and its tendency to flag all words with “me-” and “pe-” prefixes, regardless of context. These challenges highlight opportunities for future enhancements to improve the platform’s adaptability and precision.

Keywords: Damerau-Levenshtein Distance, Melting Word, RoBERTa, U-Tapis



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
KATA PENGANTAR	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi

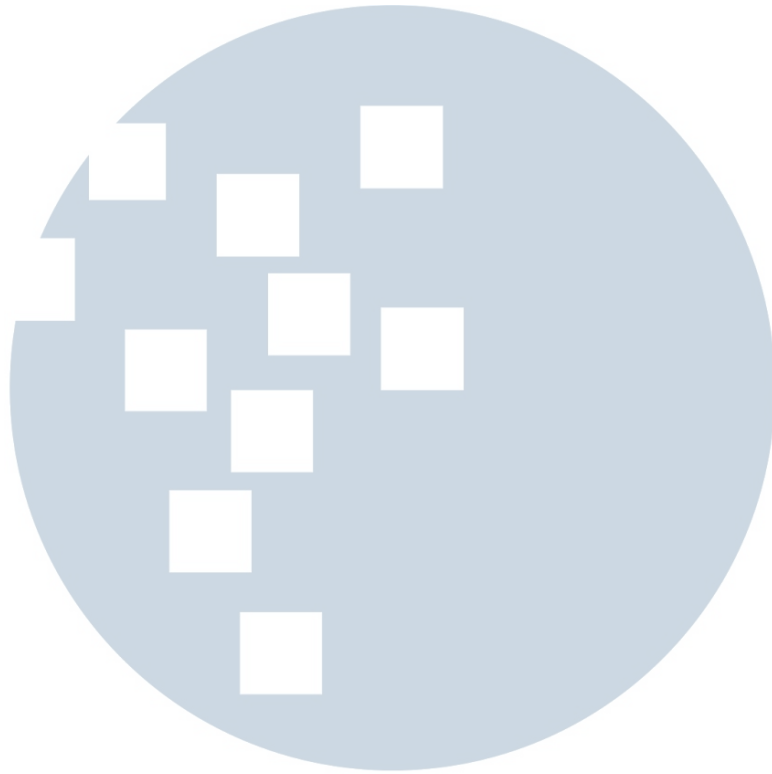


DAFTAR TABEL



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	1
Lampiran 2	Formulir Bimbingan	5
Lampiran 3	Letter of Acceptance	6
Lampiran 4	Sertifikat Partisipasi Conference	7

