

**DETEKSI KESALAHAN DAN PENGOREKSIAN PENULISAN
DALAM BAHASA JAWA MENGGUNAKAN ALGORITMA
LEVENSHTTEIN DISTANCE**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Gelar Strata Satu

Program Studi Informatika



Disusun Oleh:

Hanifa Yustikaningrum

NIM. M0514020

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SEBELAS MARET**

2021

**HALAMAN PERSETUJUAN
SKRIPSI**

**DETEKSI KESALAHAN DAN PENGOREKSIAN PENULISAN DALAM BAHASA
JAWA MENGGUNAKAN ALGORITMA LEVENSHTAIN DISTANCE**

Disusun Oleh:

**HANIFA YUSTIKANINGRUM
M0514020**

**Skripsi ini telah disetujui untuk dipertahankan di hadapan dewan penguji
pada tanggal, 28 Juli 2021**

Pembimbing I



SARI WIDYA SIHWI, S.Kom., M.T.I.
NIP. 198304122009122003

Pembimbing II



Dr. Techn. DEWI WISNU WARDANI,
S.Kom., M.S.
NIP. 197810262005012002

HALAMAN PENGESAHAN
SKRIPSI

Deteksi Kesalahan dan Pengoreksian Penulisan dalam Bahasa Jawa Menggunakan
Algoritma Levenshtein Distance

Disusun Oleh:

HANIFA YUSTIKANINGRUM
M0514020

Telah dipertahankan di hadapan dewan penguji
pada tanggal 30 Juli 2021

Susunan Dewan penguji

1. SARI WIDYA SIHWI, S.Kom., M.T.I.

NIP. 198304122009122003

()

2. Dr. Techn. DEWI WISNU WARDANI, S.Kom., M.S.

NIP. 197810262005012002

()

3. ESTI SURYANI, S.Si., M.Kom.

NIP. 197611292008122001

()

4. ARDHI WIJAYANTO, S.Kom., M.Cs.

NIP. 198806082018031001

()

Disahkan oleh

Kepala Program Studi Informatika



Dr. WIHARTO, S.T., M.Kom.

NIP. 197502102008011005

DETEKSI KESALAHAN DAN PENGOREKSIAN PENULISAN DALAM BAHASA JAWA MENGGUNAKAN ALGORITMA LEVENSHTTEIN DISTANCE

HANIFA YUSTIKANINGRUM

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Sebelas Maret

ABSTRAK

Penulisan ejaan yang benar merupakan salah satu hal yang paling penting atau komponen inti dalam sebuah artikel. Apabila terdapat penulisan ejaan yang salah atau tidak sengaja salah akibat *typo* dalam sebuah artikel maka akan mengakibatkan suatu kalimat memiliki makna yang berbeda sehingga dapat menimbulkan multitafsir bagi pembaca. Maka dari itu, diperlukan suatu pendeteksian dan proses pengoreksian sebuah artikel dalam Bahasa Jawa untuk dapat melakukan pembetulan serta koreksi apabila terdapat kesalahan penulisan dalam Bahasa Jawa. Hal ini dikarenakan masih sedikit resource yang dapat digunakan dalam menentukan deteksi kesalahan dan pengoreksian dalam Bahasa Jawa. Metode yang digunakan dalam penelitian ini adalah menggunakan dictionary lookup dan Levenshtein Distance. Penelitian yang dilakukan menghasilkan akurasi sebesar 92% dalam akurasi, 0,75 presisi, dan 0,68 untuk recall.

Kata kunci: Bahasa Jawa, Levenshtein Distance, *correction*

DETEKSI KESALAHAN DAN PENGOREKSIAN PENULISAN DALAM BAHASA JAWA MENGGUNAKAN ALGORITMA LEVENSHTein DISTANCE

HANIFA YUSTIKANINGRUM

Department of Informatics, Faculty of Mathematic and Natural Sciences,

Universitas Sebelas Maret

ABSTRACT

Writing comprehension is one of the most important or core components in an article. If there is a spelling that is wrong or accidentally wrong due to a typo in an article, it will result in a sentence having a different meaning so that it can cause different meanings for the reader. Therefore, we need a detection and correction process for an article in Javanese to be able to make corrections and corrections if there are writing errors in Javanese. This is because there are still few resources that can be used in determining error detection and correction in the Java language. The method used in this study is to use a dictionary lookup and Levenshtein Distance. The research carried out resulted in an accuracy of 92% in accuracy, 0.75 precision, and 0.68 for recall.

Key words: Bahasa Jawa, Levenshtein Distance, *correction*

MOTTO

“Sesungguhnya setelah kesulitan itu ada kemudahan”

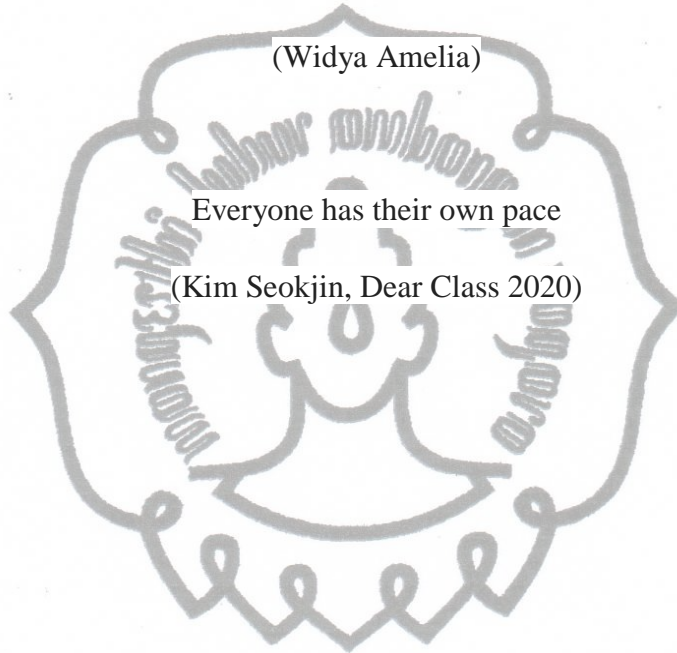
(QS. Al Insyirah ayat 5-6)

The best things is yet to come

(Widya Amelia)

Everyone has their own pace

(Kim Seokjin, Dear Class 2020)



PERSEMBAHAN

Skripsi ini saya persembahkan kepada:

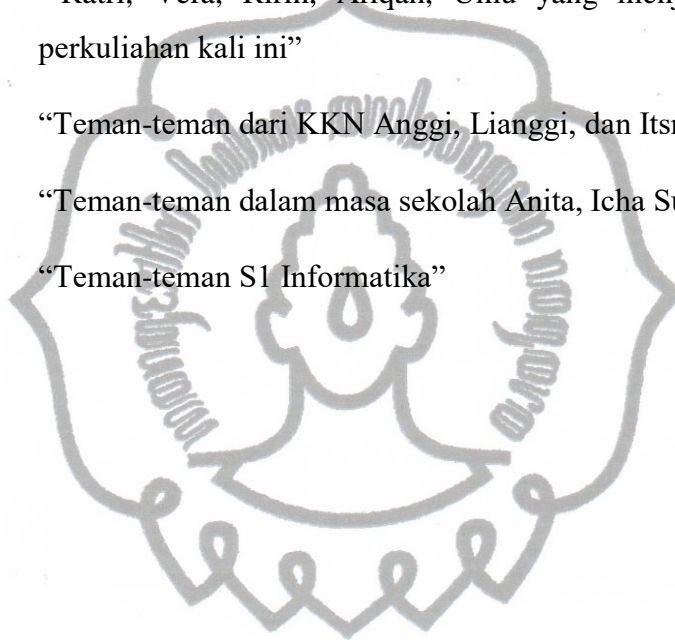
“Bapak dan Ibu serta Simbah yang telah memberikan motivasi, dukungan dalam seluruh proses kali ini”

“Ratri, Vera, Ririh, Afiqah, Umu yang menjadi dalam masa perkuliahan kali ini”

“Teman-teman dari KKN Anggi, Lianggi, dan Itsnaini”

“Teman-teman dalam masa sekolah Anita, Icha Suci, Ansheila”

“Teman-teman S1 Informatika”



KATA PENGANTAR

Puji syukur kehadirat Allah Subhanahu Wa Ta'ala atas limpahan berkah dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul Deteksi Kesalahan Penulisan Dalam Bahasa Jawa Menggunakan Algoritma Levenshtein Distance

Rasa terima kasih penulis sampaikan kepada :

1. Bapak Dr. Wiharto, S.T., M.Kom selaku Kepala Program Studi Informatika yang telah memberikan dukungan selama proses penyusunan Tugas Akhir.
2. Bapak Drs. Bambang Harjito M.App.Sc.,Ph.D. selaku Pembimbing akademik yang telah memberikan dukungan selama proses penyusunan Tugas Akhir.
3. Ibu Denis Eka Cahyani, S.Kom.,M.Kom., Ibu Sari Widya Sihwi, S.Kom., M.T.I., M.Sc., dan Ibu Dr.tech. Dewi Wisnu Wardani, S.Kom., M.S selaku Dosen Pembimbing atas ilmu, dan masukan yang diberikan, serta kesediaan untuk membimbing penulis selama pelaksanaan Tugas Akhir.
4. Bapak, Ibu, dan Simbah yang selalu senantiasa memberikan dukungan dan motivasi kepada penulis untuk menyelesaikan tugas akhir
5. Teman-teman Informatika 2014 yang telah kebersamai penulis selama di bangku perkuliahan.

Penyusunan skripsi ini tentunya masih banyak kekurangan. Oleh sebab itu, kritik dan saran pembaca sangat diperlukan. Semoga skripsi ini bisa bermanfaat bagi pembaca maupun penulis sendiri.

Surakarta, 2021

Penulis

DAFTAR ISI

SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
ABSTRAK.....	iv
ABSTRACT.....	v
MOTTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	2
BAB I.....	3
1.1 Latar Belakang.....	3
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.4 Sistematika Penulisan.....	6
BAB II.....	7
2.1 Landasan Teori.....	7
2.1.1 Levenshtein Distance.....	7
2.1.2 Spelling Error.....	8
2.1.3 Spelling Checking.....	9
2.1.4 Dictionary Lookup.....	9
2.1.5 Bahasa Jawa.....	10
2.2 Penelitian Terkait.....	11
2.3 Kerangka Pemikiran.....	15
BAB III.....	16
3.1 Pengumpulan Data.....	16
3.2 Text Preprocessing.....	17

3.3	Dictionary Lookup.....	18
3.4	Levenshtein Distance.....	18
3.5	Melakukan Evaluasi Model.....	18
BAB IV	20
4.1	Pengumpulan Data.....	20
4.2	Text Preprocessing.....	20
4.3	Fitur Dictionary Lookup.....	21
4.4	Proses perhitungan dengan Levenshtein Distance	22
4.5	Evaluasi	24
BAB V	27
5.1	Kesimpulan.....	27
5.2	Saran.....	27
DAFTAR PUSTAKA	28
LAMPIRAN	30

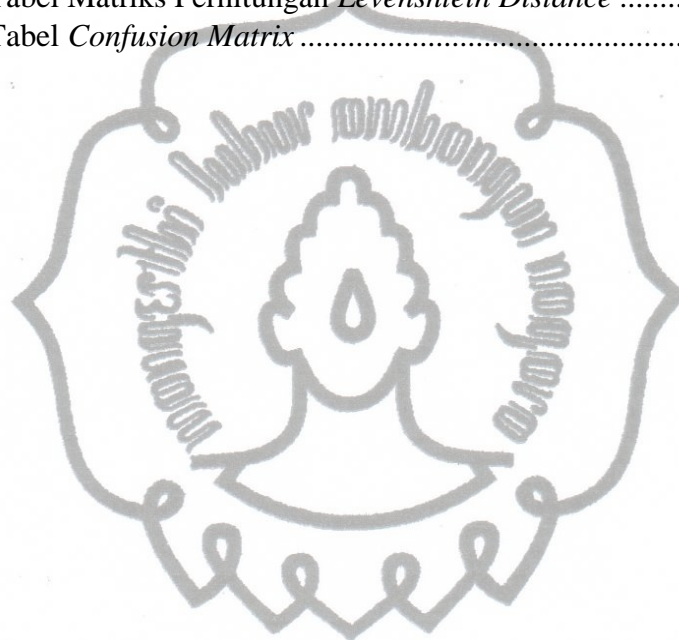
DAFTAR GAMBAR

Gambar 3. 1 Metodologi Penelitian.....	16
Gambar 3. 2 Tahapan Proses <i>Preprocessing</i>	17
Gambar 4. 1 Contoh Tahap <i>Preprocessing Case Folding</i>	20
Gambar 4. 2 Contoh Tahap <i>Preprocessing Tokenisasi</i>	21



DAFTAR TABEL

Tabel 2. 1 Tabel Perhitungan <i>Levenshtein Distance</i>	8
Tabel 2. 2 Penelitian Terkait	12
Tabel 2. 3 Penelitian Terkait Lanjutan	13
Tabel 2. 4 Penelitian Terkait Lanjutan	14
Tabel 3. 1 Tabel <i>Confusion Matrix</i>	19
Tabel 4. 1 Sampel Data Artikel Bahasa Jawa.....	21
Tabel 4. 2 Tabel Matriks Perhitungan <i>Levenshtein Distance</i>	23
Tabel 4. 3 Tabel <i>Confusion Matrix</i>	25



DAFTAR LAMPIRAN

Lampiran 1: Sample Data Penelitian.....	30
Lampiran 2: Kamus Bahasa Jawa	33

