

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**NORMALISASI TEKS BAHASA MAKASSAR MENGGUNAKAN
DAMERAU LEVENSHTein****TUGAS AKHIR**

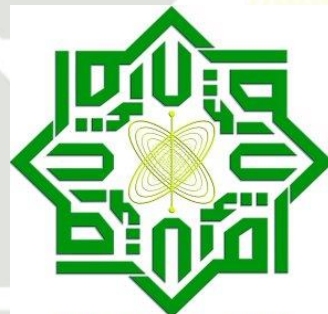
Diajukan Sebagai Salah Satu Syarat

Untuk Memperoleh Gelar Sarjana Teknik

Pada Jurusan Teknik Informatika

Oleh

MUJIBURAHMAN
11451104625



UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU

PEKANBARU

UIN SUSKA RIAU
2022

LEMBAR PERSETUJUAN

NORMALISASI TEKS BAHASA MAKASSAR MENGGUNAKAN *DAMERAU LEVENSHTTEIN*

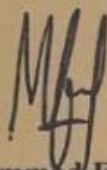
TUGAS AKHIR

Oleh

MUJIBURAHMAN
NIM. 1145110465

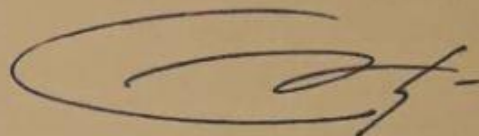
Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 11 Januari 2022

Pembimbing I,



Muhammad Fikry, ST, M.Sc
NIP. 9801018 2007101 002

Pembimbing II



Yusra, ST, MT
NIP198401232015032001

LEMBAR PENGESAHAN

NORMALISASI TEKS BAHASA MAKASSAR MENGGUNAKAN DAMERAU LEVENSHTein

Oleh

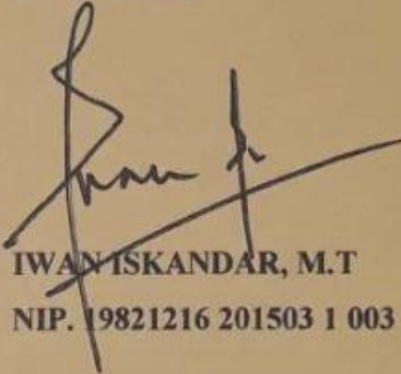
MUJIBURAHMAN
NIM. 11451104625

Telah dipertahankan di depan sidang dewan penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 11 Januari 2022

Mengesahkan,

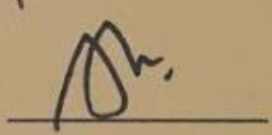
Ketua Jurusan,



IWAN ISKANDAR, M.T
NIP. 19821216 201503 1 003

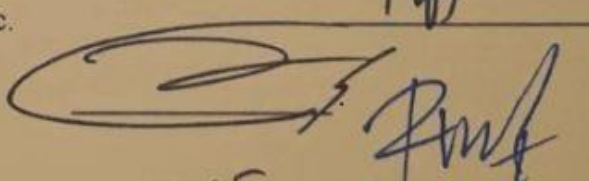

Dekan,

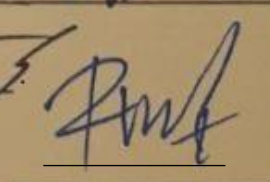
Dr. HARTONO, M.Pd
NIP. 19640301 199203 1 003

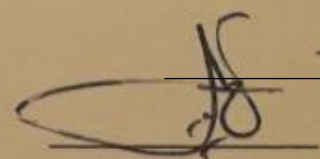
DEWAN PENGUJI

Ketua : Dr. Alwis Nazir, M.Kom 

Pembimbing I : Muhammad Fikry, ST, M.Sc. 

Pembimbing II : Yusra, ST, MT 

Penguji I : Reski Mai Candra, ST, MT 

Penguji II : Feby Yanto, M.Kom 

Lampiran Surat:

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : MUJIBURAHMAN

NIM : 11451104625

Tempat/Tgl. Lahir : Bekawan / 30 Juni 1996

Fakultas/Pascasarjana : SAINS DAN TEKNOLOGI

Prodi : TEKNIK INFORMATIKA

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:

NORMALISASI TEKS BAHASA MAKASSAR MENGGUNAKAN DAMERAU
LEVENSHTEIN

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak mana pun juga.

Pekanbaru, 20 Januari 2022

Yang membuat pernyataan



MUJIBURAHMAN
NIM. 11451104625

**pilih salah satu sesuai jenis karya tulis*



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir ini yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir 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 terdapat dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, November 2021

Yang membuat pernyataan,

MUJIBURAHMAN
11451104625

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PESEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Allah akan meninggikan derajat orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu pengetahuan.

Alhamdulillah...

Rasa syukur kuhaturkan kepada-Mu, Yaa Allah yang Maha Ber-Ilmu, hanya karena karuniaMu sajalah hamba-Mu akhirnya dapat menyelesaikan Tugas Akhir ini

Tugas Akhir ini kupersembahkan untuk Ayah dan Ibu.

Tentulah tidak akan tergantikan semua jasa, pengorbanan, tetes keringat dan rasa letih itu, hanya dengan karya kecil dariku ini... Namun semoga dengan ini, aku dapat mengukir sebaris senyum bahagia di hati Ayah dan Ibu... Jika boleh kujabarkan cinta,, tentulah tidak pernah dapat seindah rasa syukurku menjadi anakmu... Terima kasih untuk semua rangkaian do'a, kasih sayang serta ilmu yang berharga...

Dan tidak lupa kupersembahkan untuk semua adik-adikku tersayang, terima kasih untuk semua dukungan dan doa yang telah diberikan selama ini...

Juga, kupersembahkan untuk semua keluargaku dan kerabat... Semua kesulitan seolah lenyap saat mengingat bahwa aku memiliki dukungan darimu semua. Aku tahu, engkau semua berjuang jauh lebih keras dariku, namun selalu memiliki energi hebat untuk menyemangatiku...

Alhamdulillah, Allah menganugerahiku keluarga yang inda



ABSTRAK

Etnik Makassar memiliki bahasa sendiri yang disebut bahasa Makassar. Fonem bahasa Makassar sebanyak 23 buah (18 konsonan dan 5 vokal), Seringnya terjadi kesalahan dalam penulisan (*Typo*) yang menyebabkan kata-kata tersebut sulit dimengerti, tidak dimengerti atau bahkan terjadi perubahan makna dari kata tersebut. Berdasarkan masalah di atas, maka diperlukan normalisasi supaya memperbaiki kesalahan penulisan bahasa Makassar sehingga mudah dipahami dan dapat dimengerti oleh pembaca serta mengurangi kesalahpahaman bagi orang di luar masyarakat Makassar yang mau belajar atau baru belajar bahasa Makassar. Pada penelitian ini, konsep yang digunakan adalah metode *Damerau Levenstein Distance Stemming* yang digunakan dari penelitian Saharuddin dengan judul *Stemming* bahasa Makassar berbasis morfologi tahun 2019. Data yang diambil berasal dari komentar di *Youtube* yang berjumlah 400 data dari berbagai komentar yang menggunakan Bahasa Makassar. Hasil pengujian dengan menggunakan rumus *suggestion adequacy* yang didapatkan pada penelitian ini memiliki rata-rata sebesar 72,22%.

Kata kunci: Normalisasi, *Damerau Levenstein*, Bahasa Makassar.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ABSTRACT

Makassar ethnic has their own language called Makassar language. Language phonemes Makasar as many as 23 pieces (18 consonants and 5 vowels), Frequent errors in writing (typo) that causes the words to be difficult to understand, not understood or there is even a change in the meaning of the word. Based on the problem above, then normalization is needed in order to correct the writing errors of the Makassar language so that easy to understand and can be understood by readers and reduce misunderstandings for people outside the Makassar community who want to learn or are just learning the Makassar language. The stemming used is from Saharuddin's research with the title Makassar language stemming morphology-based 2019. The data taken comes from comments on youtube that totaling 400 data from various comments using Makassar language. Results testing using the suggestion adequacy formula obtained in the study This has an average of 72.22%.

Keywords: Normalization , Damerau Levenstein, Makassar language.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR



Puji syukur kehadiran Allah SWT, karena berkat rahmat dan karunia-Nya penulis mampu menyelesaikan Tugas Akhir yang berjudul “NORMALISASI TEKS BAHASA MAKASSAR MENGGUNAKAN DAMERAU LEVENSHTTEIN”. Laporan ini disusun sebagai salah satu syarat untuk meraih gelar sarjana dari Universitas Islam Negeri Sultan Syarif Kasim Riau.

Selama pelaksanaan Tugas Akhir ini, penulis banyak mendapat pengetahuan, bimbingan, dukungan, arahan serta masukan dari semua pihak yang telah membantu hingga penulisan laporan ini dapat diselesaikan. Untuk itu pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. Khairunas Rajab, M.Ag, selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Iwan Iskandar, M.T. selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Fadhilah Syafria, ST, M.Kom, CIBIA selaku Pembimbing Akademik dan Koordinator Tugas Akhir Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
5. Bapak Muhammad Fikry, ST, M.Sc, selaku Dosen Pembimbing I Tugas Akhir yang telah membantu dan memberikan banyak masukan serta saran dalam menyelesaikan Tugas Akhir.
6. Ibu Yusra, S.T., M.T, selaku Dosen Pembimbing II Tugas Akhir yang telah membantu dan memberikan banyak masukan serta saran dalam menyelesaikan Tugas Akhir.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

7. Bapak Muhammad Reski Mai Candra, ST, M.Sc, selaku penguji I Tugas Akhir yang telah memberikan banyak masukan serta saran kepada penulis.
8. Bapak Febi Yanto, M.Kom, selaku penguji II Tugas Akhir yang telah memberikan banyak masukan seta saran kepada penulis.
9. Seluruh Dosen yang mengajar pada jurusan Teknik Informatika UIN Suska Riau.
10. Kedua orang tua penulis, Ayahanda Denni dan Ibunda Siti Hajerah yang selalu menjadi sosok penyemangat dan tidak pernah berhenti berdoa untuk penulis dalam menyelesaikan Tugas Akhir ini.

Semoga laporan Tugas Akhir ini dapat bermanfaat bagi penulis dan juga untuk para pembaca laporan ini. Penulis berharap mendapatkan kritik dan saran guna memperbaiki serta pengembangan dari laporan ini ke depannya. Kritik dan saran tersebut dapat dikirim ke email penulis yaitu mujiburahman@students.uin-suska.ac.id. Akhir kata penulis ucapkan terima kasih dan selamat membaca.

Pekanbaru, November 2021

Penulis

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR ISI

LEMBAR PERSETUJUAN.....	i
LEMBAR PENGESAHAN	ii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iii
LEMBAR PERNYATAAN	v
LEMBAR PESEMBAHAN	vi
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
BAB 2 KAJIAN PUSTAKA.....	5
2.1 <i>Natural Language Processing (NLP)</i>	5
2.2 <i>Text Preprocessing</i>	6
2.3 Normalisasi.....	6
2.4 Metode <i>Damerau-Levenshtein</i>	7
2.5 Bahasa Makassar	9



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.6	Tata Bahasa makassar (Morfologi)	9
2.6.1	Prefiks.....	10
2.6.2	Sufiks.....	11
2.6.3	Infiks.....	12
2.7	Penelitian Terkait.....	13
BAB 3 METODOLOGI PENELITIAN.....		18
3.1	Identifikasi Masalah	18
3.2	Studi Literatur.....	19
3.3	Pengumpulan Data.....	19
3.4	Analisa.....	19
3.4.1	Analisa Normalisasi	19
3.5	Perancangan.....	20
3.5.1	<i>Flowchart</i>	20
3.5.2	<i>Pseudocode</i>	20
3.5.3	Perancangan <i>Database</i>	20
3.5.4	Perancangan <i>Interface</i>	20
3.6	Implementasi	21
3.7	Pengujian	21
3.8	Kesimpulan dan Saran.....	21
BAB 4 PEMBAHASAN.....		22
4.1	Analisa.....	22
4.1.1	Analisa Masalah	22
4.1.2	Analisa Kebutuhan Data.....	22
4.1.3	<i>Preprocessing</i>	27
4.1.4	Normalisasi.....	28



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.2	Perancangan.....	34
4.2.1	Perancangan <i>Flowchart</i>	34
4.2.2	Perancangan <i>Pseudocode</i>	35
4.2.3	Perancangan Basis Data	39
4.2.4	Perancangan Antar Muka	40
4.3	Implementasi	44
4.4	Batas implementasi.....	45
4.5	Lingkungan implementasi	45
4.6	Implementasi halaman antarmuka.....	45
4.6.1	Halaman Dashboard	45
4.6.2	Halaman Kamus	46
4.6.3	Halaman Data Uji.....	46
4.6.4	Halaman <i>Cleaning</i>	47
4.6.5	Halaman Normalisasi Data Uji	48
4.6.6	Halaman Perhitungan	48
4.7	Pengujian	49
4.7.1	<i>White Box</i>	49
4.7.2	Kata Uji dan Pengujian	51
5	PENUTUP	53
5.1	Kesimpulan.....	53
5.2	Saran.....	53
	DAFTAR PUSTAKA	54
	LAMPIRAN A	57
	LAMPIRAN B	58

DAFTAR GAMBAR

Gambar 2.1 Buku Tata Bahasa dan Kamus Bahasa Makassar.....	10
Gambar 3.1 Metodologi Penelitian	188
Gambar 4.1 Flowchart Pengambilan Data Dari Youtube	23
Gambar 4.2 Tampilan Exportcomments.com	224
Gambar 4.3 Tampilan Konten Youtube Makassar.....	24
Gambar 4.4 Tampilan Exportcomments.com setelah link youtube dimasukkan...24	
Gambar 4.5 Tampilan Proses Save Excel File	25
Gambar 4.6 Tampilan Data Komentar dalam Microsoft Excel	226
Gambar 4.1 Flowchart analisa sistem	34
Gambar 4.8 Perancangan Pseudocode	35
Gambar 4.9 Rancangan Antarmuka Halaman Utama	37
Gambar 4.10 Rancangan Antarmuka Halaman Kamus	37
Gambar 4.11 Rancangan Antarmuka Halaman Data Uji.....	38
Gambar 4.12 Rancangan Antarmuka Halaman <i>Cleaning</i>	39
Gambar 4.13 Rancangan Antarmuka Halaman Normalisasi Data Uji.....	39
Gambar 4.14 Rancangan Antarmuka Halaman Perhitungan	40
Gambar 4.15 Rancangan Antarmuka Halaman Kesimpulan Normalisasi	41
Gambar 4.16 Tampilan Halaman Dashboard.....	42
Gambar 4.17 Tampilan Halaman Kamus.....	43
Gambar 4.18 Tampilan Halaman Data Uji	43
Gambar 4.19 Tampilan Halaman <i>Cleaning</i>	44
Gambar 4.20 Tampilan Halaman Normalisasi Data Uji	44
Gambar 4.21 Tampilan Halaman Perhitungan.....	45
Gambar 4.22 Tampilan Halaman Hasil Perhitungan	45

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	13
Tabel 4.1 Dataset.....	26
Tabel 4.2 Proses <i>Cleaning</i>	26
Tabel 4.3 Proses <i>Case Folding</i>	26
Tabel 4.4 Proses Tokenizing	27
Tabel 4.5 Proses <i>Stemming</i>	27
Tabel 4.6 Matrik Proses Damerau.....	28
Tabel 4.7 Matrik Hasil Proses Perhitungan Damerau	30
Tabel 4.8 Nilai Damerau	30
Tabel 4.9 Matrik Proses Damerau.....	31
Tabel 4.10 Matrik Hasil Proses Perhitungan Damerau.....	33
Tabel 4.11 Nilai Damerau	33
Tabel 4.12 Data Uji.....	36
Tabel 4.13 Kata Dasar.....	36
Tabel 4.14 Hasil Pengujian <i>Cleaning</i>	46
Tabel 4.15 Hasil Pengujian <i>Case Folding</i>	46
Tabel 4.16 Hasil Pengujian <i>Tokenizing</i>	46
Tabel 4.17 Hasil Pengujian <i>Stemming</i>	47
Tabel 4.18 Hasil Kata Uji	47

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Bahasa adalah sebuah kaidah dan fungsi yang dilakukan seseorang menggambarkan kesemestaan saat berpikir. Jika seseorang dapat memahami sebuah bentuk bahasa dan fungsinya, kemudian hal itu dapat menuntunnya dalam mengungkapkan bahasa dan mengerti bahasa, berarti begitulah gambaran bagaimana caranya berpikir. Jadi, memahami sebuah bentuk kata dan kaidah atau struktur bahasa menuntun cara berpikir seseorang dan selanjutnya akan menunjukkan bagaimana seseorang tersebut mengungkapkan dan mengerti bahasa tersebut. Bahasa yang ada di dunia ini tidaklah sama. Di setiap negara memiliki bahasa yang menjadi identitas negara itu sendiri. Terkadang negara tersebut menamai bahasa sesuai dengan nama negara mereka masing-masing (Irwansyah, 2017). Hal yang sering terjadi dalam kontak sosial yang menggunakan bahasa adalah terjadinya gejala *multi* kebahasaan dan mengakibatkan timbulnya gejala peminjaman, interferensi, lahirnya bahasa baru, dan sampai kepada kepunahan bahasa (Darwis, 2011). Contoh bahasa daerah yang ada di Indonesia yaitu bahasa Makassar, Jawa, bahasa Batak, bahasa Melayu, bahasa Bali, bahasa Madura, bahasa Lampung dan lain sebagainya.

Indonesia merupakan negara yang kaya akan budaya. Oleh karena itu, masyarakatnya bukan hanya mempergunakan satu bahasa, melainkan paling sedikit dua bahasa yaitu Bahasa Indonesia dan Bahasa Daerah yang dipergunakan baik dalam lingkungan masyarakat maupun dalam lingkungan berkeluarga. Dalam lingkungan itu, bahasa merupakan alat komunikasi yang paling penting. Bahasa digunakan oleh setiap anggota masyarakat yang normal dalam kehidupan sehari-hari untuk berkomunikasi dengan anggota masyarakat yang lain. Mungkin saja suatu masyarakat belum mengenal tulisan, tetapi itu tidak berarti bahwa masyarakat itu tidak memiliki bahasa. Masyarakat itu tetap memiliki bahasa, yaitu bahasa lisan. Bahasa lisan itu merupakan bahasa alami yang tumbuh dan berkembang dalam masyarakat tersebut.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Menurut Kamus Besar Bahasa Indonesia (KBBI) bahasa adalah lambang bunyi yang tidak tetap dan sering berubah-ubah, yang digunakan oleh masyarakat untuk saling berinteraksi, bekerja sama dan mengenali diri (Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2016). Bahasa selalu berkembang dan mengalami perubahan berupa penambahan kata dengan dinamika kehidupan. Bahasa akan mengalami kepunahan jika tidak digunakan dalam bentuk media komunikasi atau objek penelitian ilmiah. Keberadaan bahasa dalam media ilmu pengetahuan akan terwujud apabila bahasa tersebut tidak kehilangan karakternya (Rosyid, 2013).

Adanya perbedaan bahasa daerah yang dimiliki antara kelompok penutur bahasa daerah yang satu dengan kelompok penutur bahasa lain menyebabkan timbulnya kesulitan dalam berkomunikasi. Untuk mengatasi hal ini, Bahasa Indonesia sebagai bahasa nasional merupakan jalan terbaik karena Bahasa Indonesia telah menarik semua etnis yang ada di Indonesia untuk digunakan dalam berkomunikasi. Situasi seperti ini dapat menciptakan kontak bahasa pada penutur yang pada mulanya hanya mengenal bahasa ibu, menjadi seorang bilingual atau dwibahasawan dengan adanya bahasa Indonesia sebagai bahasa kedua. Fenomena ini dapat ditemukan pada mayoritas masyarakat Indonesia, khususnya masyarakat dengan nuansa etnik yang lekat. Pada tahun 2019 lembaga bahasa dunia *Ethnologue: Languages of the World* mencatat jumlah bahasa daerah yang terdaftar untuk Indonesia adalah 719. Indonesia merupakan bangsa yang memiliki bahasa daerah terbesar kedua di dunia setelah Papua Nugini yang mempunyai 851 bahasa etnis. Dari jumlah tersebut, 707 adalah bahasa yang hidup dan 12 yang telah punah. Dari bahasa yang hidup, 701 adalah asli dan 6 adalah non-asli. Selain itu, 18 bersifat institusional, 73 sedang berkembang, 188 kuat, 347 dalam kesulitan, dan 81 sedang sekarat (Ethnologue, 2019).

Badan Pengembangan dan Pembinaan Bahasa (Badan bahasa) juga telah memetakan bahasa daerah yang tersebar di Indonesia pada tahun 2018 sebanyak 668 bahasa. Angka bertambah setelah pada tahun 2017 hanya mencatat 625 bahasa daerah. Hasil pencatatan bahasa daerah Indonesia tersebut dilakukan sejak tahun 1992. Jumlah yang dipetakan oleh Badan Bahasa kini masih berpeluang untuk bertambah atau berkurang terhadap penelitian sebelumnya. Badan bahasa

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menganalisis dan menghimpun kata tersebut berdasarkan faktor-faktor seperti jumlah bahasa daerah, wilayah pemakai, jumlah penuturnya dan status kebugaran dari bahasanya (CNN Indonesia, 2018).

Etnik Makassar memiliki bahasa sendiri yang disebut bahasa Makassar. Aksaranya disebut aksara lontarak jangang-jangang (mirip dengan aksara Jawa/Kawi, sedangkan aksara Bugis disebut aksara bilang-bilang/sulapak appak mirip dengan aksara Rejang (Enre dalam Manyambeang, 1996: 30). Fonem bahasa Makassar sebanyak 23 buah (18 konsonan dan 5 vokal), sedangkan fonem bahasa Bugis sebanyak 29 buah (22 konsonan dan 7 vokal). Seringnya terjadi kesalahan dalam penulisan (*Typo*) yang menyebabkan kata-kata tersebut sulit dimengerti, tidak dimengerti atau bahkan terjadi perubahan makna dari kata tersebut. Beberapa contoh kalimat yang terdapat kesalahan dalam penulisan (*typo*) seperti; “Sering2ki bikin video daeng”. Terdapat kesalahan penulisan pada kata “sering2ki” yang seharusnya ditulis dengan kata “sering-sering ki”. Kesalahan penulisan (*Typo*) seperti dijelaskan di atas sering kali terjadi yang membuat pembacanya tidak mengerti kata-kata yang terdapat kesalahan penulisan (*Typo*) tersebut.

Berdasarkan masalah di atas, maka diperlukan normalisasi supaya memperbaiki kesalahan penulisan bahasa Makassar sehingga mudah dipahami dan dapat dimengerti oleh pembaca serta mengurangi kesalahpahaman bagi orang di luar masyarakat Makassar yang mau belajar atau baru belajar bahasa Makassar. Dengan Normalisasi ini juga saya sebagai peneliti dapat mengetahui bagaimana persentase keberhasilan metode normalisasi yang saya gunakan dalam penelitian ini. Normalisasi kata adalah suatu proses dalam mengembalikan sebuah bentuk kata yang tidak beraturan ke bentuk kata yang beraturan. Normalisasi kata juga berarti untuk mengembalikan kata singkatan ke bentuk kata yang sebenarnya. Proses normalisasi kata merupakan sebuah tahapan *preprocessing* data sebelum mengaplikasikan *Natural Language Processing* (NLP). Bentuk kata yang tidak beraturan tersebut misalnya seperti akronim, deretan angka, kata campuran dan sebuah singkatan. Algoritma *Damerau Levenshtein Distance* adalah algoritma pencocokan *string word to word* sebagai pengembangan dari algoritma *Levenshtein Distance*. Algoritma *Damerau Levenshtein Distance* menentukan nilai operasi



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

minimum yang dibutuhkan untuk mengubah satu *string* menjadi *string* lain, di mana operasi yang digunakan yaitu *insertion*, *deletion*, *substitution*, dan *transposition* (Damerau, 1964).

Menurut (Sutisna & Adisantoso, 2010) membahas Koreksi Ejaan *Query* Bahasa Indonesia Menggunakan *Algoritme Damerau Levenshtein*. Implementasi *Algoritme Damerau Levenshtein* untuk koreksi ejaan pada *search engine*, dapat meningkatkan kinerja temu kembali dan *query* menjadi lebih optimal. Ini terlihat peningkatan secara rata-rata *precision* sebesar 44,82% setelah dilakukan pengoreksian. Penelitian (Maghfira, 2017) menggunakan algoritma *Damerau Levenshtein Distance* mengatakan bahwa algoritma *Damerau Levenshtein Distance* menghasilkan kandidat kata dengan nilai jarak edit operasi terkecil. Pada penelitian ini, konsep yang digunakan adalah metode *Damerau Levenshtein Distance*.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, yang menjadi rumusan masalah pada penelitian Tugas Akhir ini “Bagaimana melakukan normalisasi teks berbahasa Makassar menggunakan metode *Damerau Levenshtein Distance*”.

1.3 Batasan Masalah

Supaya pembahasan masalah dapat lebih terarah dengan baik, maka harus dilakukan batasan masalah yang akan dibahas, yaitu:

1. *Stemming* yang digunakan dari penelitian Saharuddin dengan judul *Stemming* bahasa Makassar berbasis morfologi tahun 2019.
2. Data yang diambil berasal dari komentar di Youtube yang berjumlah 400 data dari berbagai komentar yang menggunakan Bahasa Makassar.

1.4 Tujuan Penelitian

Penelitian Tugas Akhir ini bertujuan untuk mengimplementasikan normalisasi teks bahasa Makassar menggunakan metode *Damerau Levenshtein Distance*.

1.5 Sistematika Penulisan

Sistematika pembahasan tugas akhir ini antara lain:

BAB 1 PENDAHULUAN



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Bab 1 menjelaskan tentang permasalahan yang akan dijadikan latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan.

BAB 2 KAJIAN PUSTAKA

Bab 2 menjelaskan tentang dasar-dasar teori yang digunakan sebagai landasan dalam pembuatan Tugas Akhir ini seperti: Pengertian Algoritma, Morfologi bahasa Makassar, pengertian Normalisasi dan penelitian terkait.

BAB 3 METODOLOGI

Bab 3 merupakan uraian yang membahas tentang tahapan yang akan dilakukan dalam penelitian dan metode apa yang akan digunakan. Tahapan yang di maksud yaitu Pendahuluan, Analisa, Perancangan, Implementasi, Pengujian dan Penutup.

BAB 4 PEMBAHASAN

Bab 4 berisi uraian analisa dan langkah kerja serta perancangan Algoritma yang akan dibuat sesuai dengan metodologi yang digunakan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari penelitian yang telah dilaksanakan dan saran perbaikan untuk kedepannya.

BAB 2

KAJIAN PUSTAKA

2.1 *Natural Language Processing (NLP)*

Natural Language Processing (NLP) dapat didefinisikan sebagai kemampuan komputer untuk memproses suatu bahasa, baik lisan maupun tulisan untuk percakapan sehari-hari yang digunakan oleh manusia. Untuk proses komputasi, bahasa harus direpresentasikan sebagai rangkaian simbol yang memenuhi aturan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tertentu. Secara sederhana, NLP mencoba untuk membuat komputer dapat mengerti perintah-perintah yang ditulis dalam standar Bahasa manusia. (Lisangan dkk, 2013)

2.2 **Text Preprocessing**

Text preprocessing adalah sebuah proses untuk mempersiapkan data sebelum melakukan proses yang lainnya. Pra proses data dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data tersebut menjadi bentuk yang lebih mudah diproses oleh sistem. Proses *preprocessing* ini meliputi beberapa proses yaitu:

1. *Cleaning* yakni proses dalam melakukan pembersihan dari atribut-atribut yang tidak dibutuhkan dengan informasi yang ada pada data. Informasi yang akan dibersihkan berupa data yang memakai *mention*, *emoticon*, *URL*, dan *hashtag*.
2. *Case Folding* yakni sebuah proses yang mengubah huruf yang terdapat pada sebuah kata menjadi ukuran huruf kecil.
3. *Tokenizing* yakni sebuah proses tahap pemotongan *string* yang di *input* berdasarkan tiap kata yang menyusunnya. Dapat juga diartikan sebagai proses memecah sekumpulan karakter dalam sebuah kalimat ke dalam satuan kata, serta bagaimana membedakan karakter-karakter tertentu yang akan diperlakukan sebagai pemisah kata atau bukan.
4. *Filtering* yakni proses tahapan pengambilan kata-kata yang penting dari hasil *token*. Proses ini bisa menggunakan algoritma *stoplist* (membuang kata yang dianggap kurang penting) atau *wordlist* (menyimpan kata penting).
5. *Stemming* yakni suatu proses untuk memperkecil jumlah sebuah indeks yang berbeda dari sebuah dokumen, dan juga untuk melakukan pengelompokan kata-kata lain yang memiliki kata dasar lain yang mempunyai kata dasar dan arti yang sama namun mempunyai bentuk atau *form* yang berbeda karena mendapatkan sebuah imbuhan yang berbeda.

2.3 **Normalisasi**

Normalisasi kata adalah proses yang dilakukan untuk mengubah kata-kata yang tidak baku menjadi kata-kata yang baku dalam suatu kalimat. Proses normalisasi kata (Buntoro dkk, 2014), meliputi:



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Memisahkan tanda baca dan simbol selain alfabet
2. Mengubah semua huruf dalam dokumen menjadi huruf kecil
3. Mengubah kata yang tidak baku menjadi kata yang baku sesuai dengan tata bahasa
4. Menghilangkan huruf berulang atau berlebih, dan
5. Menghilangkan *emoticon*.

Normalisasi kata adalah sebuah tahapan untuk mengubah bentuk suatu kata yang tidak baku dalam suatu kalimat pada tweet menjadi suatu kata yang baku dalam suatu kalimat. Normalisasi kata berguna untuk mengatasi permasalahan seperti banyaknya penyingkatan pada sebuah kata, penggunaan bahasa gaul, terjadinya kesalahan eja pada sebuah kalimat, atau penggunaan bahasa yang masih belum sesuai dengan kamus. Dengan dilakukannya sebuah proses normalisasi kata, maka permasalahan tersebut dapat diselesaikan dan akan membuat sebuah tweet yang dapat diproses dan dapat dianalisis dengan baik (Sari & Adikara, 2019).

2.4 Metode Damerau-Levenshtein

Algoritma *damerau-levenshtein distance* umumnya adalah sebagai berikut:

- a. Inisialisasikan n sebagai panjang karakter dari s dan m sebagai panjang karakter dari t. Jika n = 0 atau m = 0, maka kembalikan nilai (*return value*) berupa jarak edit dengan rumusan: jarak_edit = max(n, m) lalu lompat ke langkah 7.
- b. Buat sebuah matriks d sebanyak m + 1 baris dan n + 1 kolom
- c. Isi baris pertama dengan 0..n dan isi kolom pertama dengan 0..m.
- d. Periksa setiap karakter dari s terhadap t. Jika s[i] = t[j] maka cost = 0. Jika s[i] ≠ t[j] maka cost = 1.
- e. Isikan nilai dari setiap sel d[i, j] baris per baris dengan : d[i, j] = min(x, y, z)

keterangan :

d[i, j] : sel yang merupakan pertemuan kolom j dengan baris i pada matriks

d.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- x : adalah nilai yang terdapat di sel atas dari posisi sel sekarang ditambah dengan 1 (satu) atau dapat dirumuskan : $x = d[i - 1, j] + 1$
- y : nilai yang terdapat di sel sebelah kiri dari posisi sel sekarang ditambah 1 (satu) atau dapat dirumuskan : $y = d[i, j - 1] + 1$
- z : nilai yang terdapat di sel sebelah atas dari sebelah kiri sel sekarang (arah barat laut) ditambah nilai cost dan dapat dirumuskan : $z = d[i - 1, j - 1] + \text{cost}$

Jika $i > 1$ dan $j > 1$ dan $s[i] = t[j - 1]$ dan $s[i - 1] = t[j]$ yang mana artinya setelah kedua kata dibandingkan terdapat karakter yang dapat ditransposisikan, maka isi nilai sel $d[i, j]$ dengan rumusan berikut: $d[i, j] = \min(d[i, j], d[i - 2, j - 2] + \text{cost})$

- f. Setelah langkah iterasi di atas selesai, maka jarak edit akan ditemukan pada sel $d[n, m]$ yaitu sel pada pojok kanan baris terakhir.
- g. Selesai Penelitian yang dilakukan oleh Iskandar Setiadi yang berjudul “Damerau-Levenshtein Algorithm and Bayes Theorem for Spell Checker Optimization” mengatakan bahwa dalam ilmu komputer dan teori informasi Algoritma *Levenshtein-distance* adalah salah satu metode untuk pengolahan string, dimana *Levenshtein distance* digunakan untuk mengukur perbedaan antara dua buah kata yang dibentuk dari susunan *string*, *Levenshtein-distance* dapat digunakan untuk mengolah perbedaan dua buah kata atau kalimat sehingga dapat disarankan kata atau kalimat mana yang paling mendekati kata yang ingin dicari, Iskandar Setiadi dalam penelitiannya berhasil menggabungkan algoritma *Damerau Levenshtein* dengan *teorama bayes* untuk menambah keakuratan dalam pengejaan kata yang di proses.

Pada algoritma *Damerau Levenshtein-distance* ini, digunakan rumus *Suggestion Adequacy (SA)* untuk mendapatkan nilai akurasinya. Untuk pemberian skor kata pada kata yang tidak normal, dilakukan dengan cara sebagai berikut:

1. Jika rekomendasi kata pertama sama dengan hasil pakar, maka bernilai 1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Jika rekomendasi kata pertama tidak sama dengan hasil pakar, namun rekomendasi kata kedua sampai seterusnya sama dengan hasil pakar, maka bernilai 0,5
3. Jika rekomendasi kata tidak sama satu pun dengan hasil, maka bernilai -0,5
4. Jika tidak memiliki rekomendasi, maka bernilai 0

Setelah dilakukan pemberian skor pada kata yang tidak normal, maka digunakan rumus *suggestion adequacy* untuk melakukan perhitungan pengujian.

$$SA = \frac{\sum_{k=0}^n SK}{N_{Tn}}$$

Keterangan:

SA: *Suggestion Adequacy*

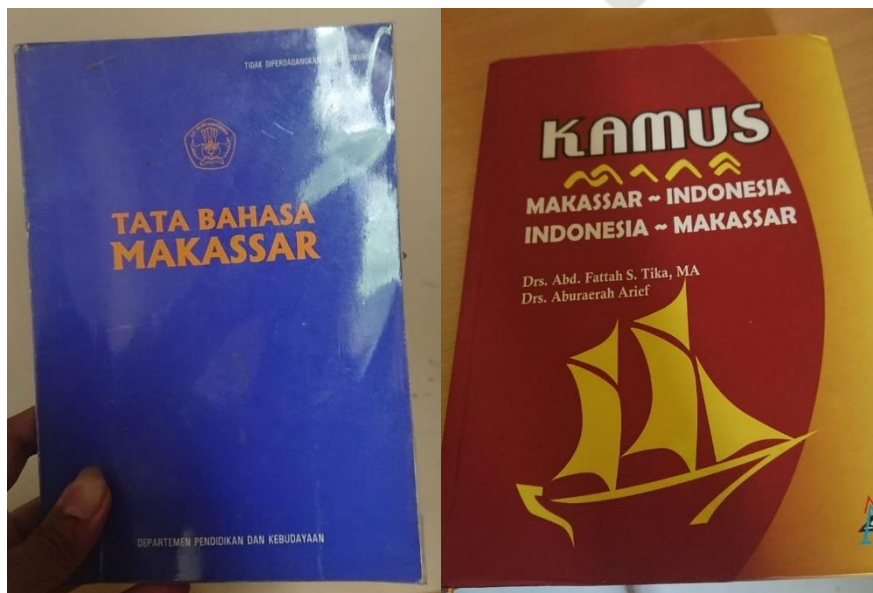
SK: jumlah seluruh skor

Tn: jumlah seluruh kata tidak normal

2.5 Bahasa Makassar

Menurut (Manyambeang dkk, 1996), bahasa Makassar yang merupakan bahasa yang hidup dan menjadi alat komunikasi masyarakat pemakainya, bahkan menjadi pendukung kebudayaan di Sulawesi Selatan, cukup luas daerah pemakainya. Bahasa ini menjadi bahasa kedua setelah bahasa Bugis di Sulawesi selatan. Bahasa Makassar dipergunakan di sebagian selatan Sulawesi Selatan.

2.6 Tata Bahasa makassar (Morfologi)



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Gambar 2.1 Buku Tata Bahasa dan Kamus Bahasa Makassar

Morfologi merupakan cakupan ilmu tentang linguistik yang mempelajari bagian-bagian kata, pembentukan kata dan penataan kata. Pengertian dari morfologi itu sendiri ialah suatu tahap pembentukan kata yang berasal dari *laksem* yakni satuan kata dan leksikal yang disebut satuan gramatikal (Maryam, 2016). Satuan morfologi adalah morfem (akar atau afiks) dan kata. Proses morfologi melibatkan komponen, antara lain: komponen dasar atau bentuk dasar, alat pembentuk dan makna gramatikal (Hestiani, 2012). Menurut Ramadhanti dkk (2019) proses morfologi ialah suatu tahap pembentukan kata yang berasal dari sebuah bentuk kata dasar. Morfologi bahasa bekerja dengan menganalisis proses morfologi itu sendiri, misalnya yaitu proses morfologi dalam pembentukan kata “membantu” merupakan kata yang memiliki kata dasar “bantu” dan ditempelkan kata imbuhan yaitu mem-. Morfologi dalam bahasa Makassar terdapat tiga kelompok imbuhan, yaitu: awalan, sisipan, dan akhiran (Manyambeang dkk, 1996)

3.51 Prefiks

1. Prefiks Ma-

Prefiks mi mengandung arti mempunyai sifat, (menjadi) seperti yang dinyatakan dalam kata dasarnya.

Contoh:

ma+eja = maeja '(me)merah'

rna+rannu marannu '(ber)gembira'

2. Prefiks Pa-

Contoh prefiks pa-

pa + inung = painung 'peminum'

pa + jama = pajama 'pekerja'



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

pa + dekdek = padedek 'pandai (tukang)'

3. Prefiks pak-

Contoh prefiks pak-

pap + parek = papparek 'alat pembuat'

pap + barrasak = pakbarrasak 'sapu (penyapu)'

pat + tongkok = pattongkok 'atap, penutup'

Sufiks

1. Sufiks — — i

Sufiks mi mengandung arti:

- a) Melakukan pekerjaan dengan menggunakan alat yang disebut dalam kata dasar. Contoh: batu+i = batui 'lempari dengan batu'
- b) 'berulang-ulang'
contoh: sambila + i = sambilai 'lempari'
- c) 'buang', 'keluarkan', 'ambil'
contoh: sissik + i = sissiki 'keluarkan sisiknya'

2. Sufiks — ang

Sufiks mi umumnya sama artinya dengan akhiran - an, dalam bahasa Indonesia. Sufiks mi mengandung arti:

- a) menyatakan 'arah'
- b) contoh: Ulu + an = uluan 'arahkan (tunjukkan jalan)'
- c) alat 'untuk'
contoh: bulek + an = bulekan) 'usungan'

3. Sufiks — ma

Akhiran -- ma mi selalu diikuti (bersambung) dengan kata ganti nama pelaku, - ak, - ko, - i. Sufiks mi mengandung arti "sudah".

Contoh:

erok+ma+ak = erokmak 'Saya sudah mau'

eroki-rna+ko = erokmako 'Kau sudah mau'

4. Sufiks — mi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Mi dapat juga dipandang sebagai satu sufiks tersendiri dengan arti sama dengan lah. dalam bahasa Indonesia. (Jadi merupakan partikel).

Contoh:

- aile + mi = allemi 'ambillah'
- erang + mi = erammi 'bawalah'
- erang + mi = erammi 'bawalah'

5. Sufiks — ka

Sufiks mi sama artinya dengan akhiran (partikel) kah dalam bahasa Indonesia.

Contoh:

- inai + ka inaika 'siapakah'
- niak + ka niakka 'adakab'

3.53 Infiks

1. Infiks — um —

Infiks mi mengandung arti melakukan pekerjaan seperti yang dinyatakan dalam kata dasar.

Contoh:

- s + um + elan = sumelan 'menyetam'
- s + um + ayak sumayak 'terbang menurun'

2. Infiks—im-

Infiks mi mengandung arti: Melakukan perjalanan dengan menumpang kendaraan yang memakai alat seperti yang disebut dalam kata dasar.

Contoh:

- s + im + ombalak = simombalak 'berlayar'

3. Infiks —al —

Infiks mi mengandung arti:'banyak sekali':

Contoh:

- g + al + akruk = galakruk 'gaduh'

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4. Infiks—ar-

Infiks mi mengandung arti: mengeraskan arti kata dasar.

Contoh:

k + ar + ankan = karanka 'genggam'

Penelitian Terkait

Beberapa penelitian terkait tentang normalisasi adalah sebagai berikut:

Tabel 2.1 Penelitian Terkait

No	Peneliti	Judul	Metode	Hasil
1	(Maghfira, 2017)	Deteksi Kesalahan Ejaan dan Penentuan Rekomendasi Koreksi Kata yang Tepat Pada Dokumen Jurnal JTIK Menggunakan <i>Dictionary Lookup</i> dan <i>Damerau-Levenshtein Distance</i>	<i>Dictionary Lookup</i> dan <i>Damerau-Levenshtein Distance</i>	Kesimpulan dari penelitian ini yaitu Metode <i>Dictionary Lookup</i> dan <i>Damerau-Levenshtein Distance</i> dapat diimplementasikan dengan baik pada proses deteksi dan koreksi kesalahan ejaan kata pada jurnal JTIK. Pada skenario pengujian jumlah kesalahan ejaan kata didapatkan nilai presisi dan recall terbaik sebesar 0.76 dan 0.99. sedangkan nilai presisi dan recall terbaik sebesar 0.78 dan 1. Berdasarkan hasil pada kedua skenario menunjukkan bahwa nilai recall lebih tinggi daripada nilai presisi. Hasil tersebut menunjukkan bahwa semua koreksi kata yang diharapkan berhasil diambil oleh sistem sebagai hasil koreksi kata. Selain itu berdasarkan hasil pengujian yang didapatkan, jumlah kesalahan ejaan kata dan jumlah kata dalam dokumen tidak terlalu berpengaruh secara signifikan terhadap kinerja koreksi sistem. Hasil koreksi kesalahan kata lebih dipengaruhi oleh kelengkapan kata pada kamus sebagai acuan kandidat koreksi kata dan tipe kesalahan ejaan kata.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti	Judul	Metode	Hasil
2	(Lusetti, 2018)	<i>Encoder-Decoder Methods for Text Normalization</i>	<i>Encoder-Decoder</i>	Hasil dari penelitian ini bahwa mengintegrasikan model bahasa tingkat yang berbeda ke dalam neural encoder-kerangka kerja decoder memungkinkan metode untuk mencapai dan bahkan meningkatkan kinerja karakter- metode terjemahan mesin statistik tingkat, yang sebelumnya dianggap lebih unggul daripada metode saraf dalam tugas normalisasi teks. Metode yang kami usulkan adalah adaptasi dari mekanisme yang diperkenalkan di Indonesia terjemahan mesin dan segmentasi morfologis. Sedangkan percobaan dilakukan dalam tulisan ini menunjukkan keuntungan dari mengintegrasikan model bahasa tingkat yang berbeda, adaptasi yang kami usulkan bisa jadi diperluas untuk mengintegrasikan skor potensial lainnya ke dalam kerangka encoder-decoder tunggal. Kemungkinan ini dapat dieksploitasi untuk perbaikan lebih lanjut dari metode normalisasi teks.
3	(Tri Sony Saragih, 2017)	Normalisasi Teks Pada Teks <i>Twitter</i> Berbahasa Indonesia Menggunakan Algoritme Jarak <i>String</i> Pada R	<i>Levenshtein distance, Hamming Distance, Soundex,</i>	Hasil penelitian ini dari perbandingan 10 metode, akurasi yang paling tinggi adalah menggunakan metode lcs yaitu 69% dan menggunakan data korpus Kompas yang diseleksi sesuai KBBI sebagai kamus. Dengan demikian, dalam implementasi fungsi ini dibuat default menggunakan metode lcs. Fungsi ini telah mampu melakukan perubahan kata tidak baku menjadi baku, meskipun perubahan yang dilakukan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti	Judul	Metode	Hasil
				tidak semuanya benar sesuai dengan yang seharusnya. Hal ini disebabkan oleh adanya saran kata lain yang memiliki jarak minimal yang sama dan tidak adanya kata dengan perbaikan seharusnya pada kamus.
4	(Hany Hasan & Arul Menezes, 2013)	<i>Social Text Normalization using Contextual Graph Random Walks</i>	<i>Contextual Graph Random Walks</i>	Hasil dari penelitian ini adalah dengan menggunakan sistem sebagai preprocess-langkah untuk sistem terjemahan mesin yang meningkatkan kualitas terjemahan sebesar 6%. Sebagai perpanjangan dari pekerjaan ini, juga akan memperluas pendekatan untuk menangani ke lebih banyak Bahasa.
5	(Joko Jatminto & I Kadek Dwi Nuryana, 2016)	<i>Implementasi Spelling Checker Dengan Algoritma Levenshtein Distance Pada Ensiklopedia It (Information Technology) Berbasis imwebsite</i>	<i>Levenshtein Distance</i>	Hasil dari penelitian ini adalah Metode levenshtein distance dapat diterapkan untuk melakukan proses pengkoreksian kata kunci pada aplikasi pencarian berbasis web. dalam jumlah yang banyak dengan durasi waktu eksekusi yang tergantung dari kompleksitas kata kunci yang di input-kan. Dapat membantu mengatasi permasalahan pada kesalahan pengetikan dengan mekanisme penambahan, penyisipan dan penghapusan karakter. Algoritma levenshtein distance akan bekerja lebih baik jika kedua data yang dibandingkan mempunyai urutan kata yang sama. Algoritma levenshtein distance akan lama dalam proses jika kedua data yang dibandingkan mempunyai urutan kata yang tidak sama dan banyak karakter yang berbeda.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti	Judul	Metode	Hasil
6	(Lisangan, 2015)	Implementasi <i>N-Gram Technique</i> Dalam Deteksi <i>Plagiarisme</i> Pada Tugas Mahasiswa	<i>N-Gram Technique</i>	Hasil dari penelitian ini adalah Semakin besar nilai <i>n</i> yang diberikan akan memberikan tingkat relevansi yang semakin baik dengan syarat nilai rata-rata selisih relevansi lebih besar dari 0 (nol). Dalam penelitian ini, nilai <i>n</i> terbaik yang diperoleh adalah 7.
7	(Sutisna & Adisantoso, 2010)	Koreksi Ejaan <i>Query</i> Bahasa Indonesia Menggunakan <i>Algoritme Damerau Levenshtein</i>	<i>Algoritme Damerau Levenshtein</i>	Implementasi algoritma yang digunakan pada sistem ini dapat meningkatkan kinerja penemuan kesalahan lebih optimal jika diimplementasikan dengan baik dan sempurna sesuai apa yang sudah diinputkan ke <i>database</i>
8	(Fahma, 2018)	Identifikasi Kesalahan Kata (<i>Typograogical Error</i>) pada Dokumen Berbahasa Indonesia Menggunakan Metode <i>N-Gram</i> dan <i>Levenshtein Distance</i>	<i>N-Gram</i> dan <i>Levenshtein Distance</i>	Hasil presisi dan recall pada penelitian ini memiliki nilai yang beragam pada setiap skenario pengujian. Nilai presisi terbaik yang dihasilkan sistem sebesar 0.97 pada skenario pengujian <i>typographical error</i> jenis <i>insertion</i> . Sementara itu, nilai recall terbaik yang dihasilkan sistem sebesar 1 pada skenario pengujian <i>typographical error</i> jenis <i>substitution</i> .
10	(Yulianto, 2018)	<i>Autocomplete and Spell Checking levenshtein Distance Algorithm to Getting Text Suggest Error Data Searching In Library</i>	<i>Levenshtein Distance</i>	Mengembangkan sistem pemeriksa ejaan harus menggunakan NLP dan membutuhkan basis data yang besar guna menyimpan semua kosa kata yang terdapat dalam kamus dan tata bahasa
11	(Simanjuntak, 2018)	<i>Spelling Corrector</i> Bahasa Indonesia dengan Kombinasi	<i>Peter Norvig</i> dan <i>N-Gram</i>	Kombinasi kedua metode ini dapat digunakan dalam memperbaiki kesalahan pengetikan, walaupun tidak dapat memperbaiki kata dengan tingkat kesalahan dua



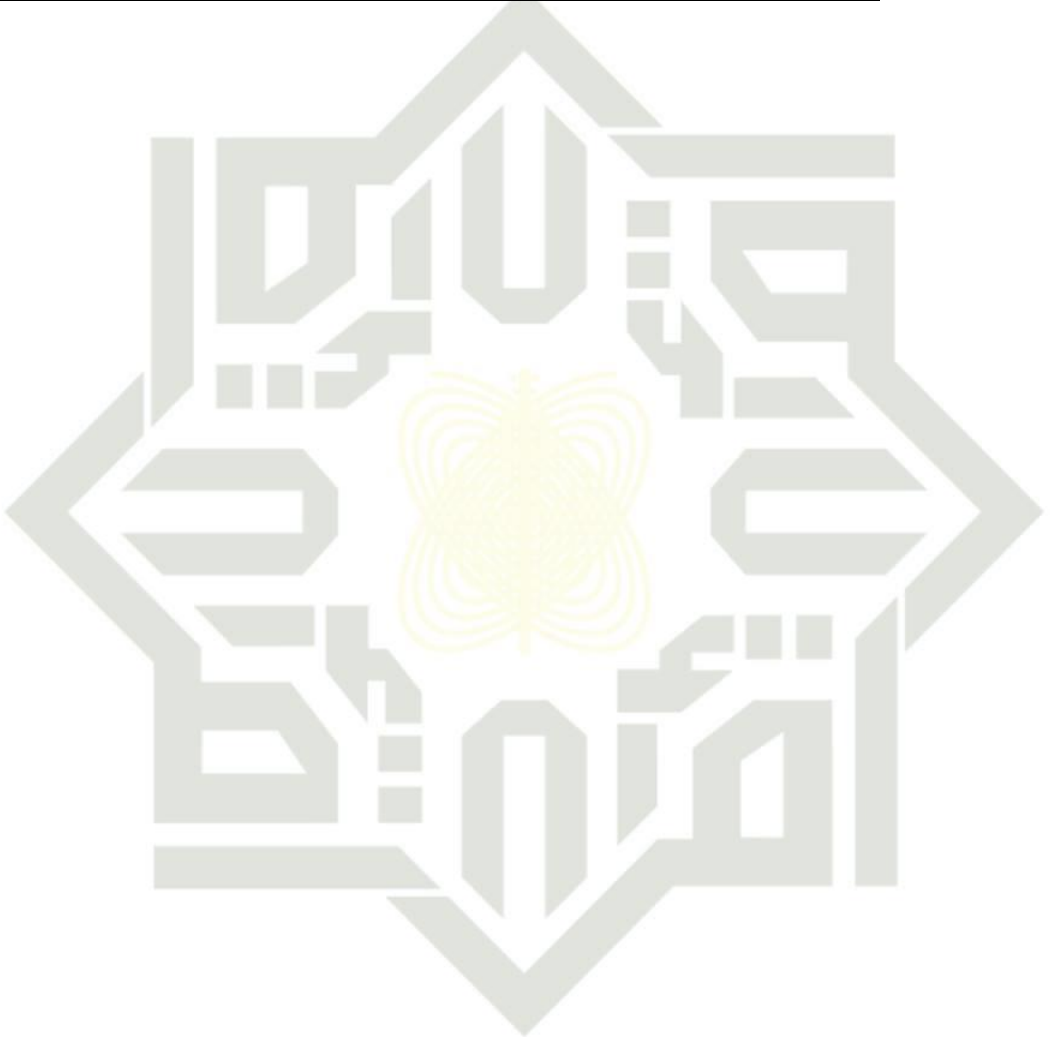
No	Peneliti	Judul	Metode	Hasil
		Metode <i>Peter Norvig</i> dan <i>N-Gram</i>		huruf atau lebih. Metode ini sangat bergantung pada korpus yang digunakan. Jika semakin baik korpus yang digunakan, maka semakin baik pula sistem menemukan kata yang relevan.

© Hak Cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

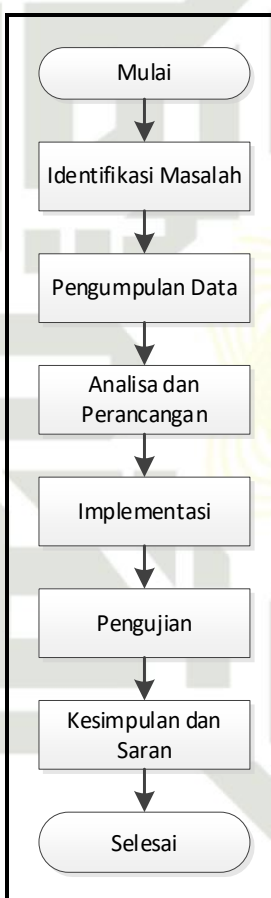


UIN SUSKA RIAU

BAB 3

METODOLOGI PENELITIAN

Metodologi penelitian adalah proses untuk menemukan hasil dari nilai pengujian dan pengembangan melalui pengetahuan, pengumpulan data dan analisa data secara sistematis. Berikut tahapan untuk melakukan penelitian pada tugas akhir dengan alur penelitian sebagai berikut:



Gambar 3.1 Metodologi Penelitian

3.1 Identifikasi Masalah

Pada tahap ini akan menjadi tahap awal dari melakukan penelitian tersebut, yang akan bertujuan untuk memahami permasalahan pada penelitian ini. Adapun inti permasalahan pada penelitian ini adalah bagaimana cara mengimplementasikan algoritma normalisasi teks berdasarkan aturan tata bahasa Makassar.

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.2 Studi Literatur

Studi literatur berguna sebagai pembelajaran teori untuk penelitian yang sedang dilakukan, karena hal tersebut untuk melandasi pengujian dan analisis yang akan dilaksanakan. Peneliti mengambil dari jurnal, skripsi serta artikel yang berhubungan dengan metode *Damerau Levensthein*. Sehingga diharapkan nantinya dapat membantu peneliti tentang analisis, teknik dan strategi dalam melaksanakan penelitian.

3.3 Pengumpulan Data

Pengumpulan Data adalah tahap yang berkaitan dengan penelitian ini dan digunakan untuk keberhasilan penelitian menggunakan data yang didapat dari kamus bahasa Makassar, morfologi bahasa Makassar, dan dari media sosial Facebook dan Youtube berupa percakapan (Komentar) yang menggunakan bahasa Makassar. Morfologi yang dikumpulkan berguna untuk analisa tahapan normalisasi dan komentar dari video Youtube digunakan untuk tahapan pengujian. Beberapa contoh komentar *typo* di Youtube dapat dilihat pada lampiran. Data berupa komentar tersebut yang akan digunakan dan dikumpulkan sebanyak 400 komentar yang menggunakan bahasa Makassar.

3.4 Analisa

Dalam tahapan ini, data-data yang telah didapatkan sebelumnya akan di analisa. Tahap ini bertujuan untuk mendapatkan suatu gambaran yang jelas tentang penelitian yang segera dilakukan. Berikut beberapa proses analisa:

3.4.1 Analisa Normalisasi

Tahap pertama akan dilakukan adalah memasukan kata yang akan di normalisasi. Selanjutnya melakukan tahap *cleaning* berupa proses pembersihan kata dari simbol yang tidak diperlukan seperti, *hashtag*, *url* dan lain-lain. Kalimat yang sudah melalui tahap *cleaning* akan melakukan proses yaitu *case folding* yang bertujuan untuk mengubah semua huruf menjadi huruf besar atau huruf kecil. Tahap selanjutnya melakukan *tokenizing* yaitu proses pemotongan kalimat di mana setiap kata terkadang memiliki satu kalimat atau lebih.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tahap selanjutnya kalimat yang sudah dipotong atau dipisah akan melalui tahap *filtering* untuk mengambil kata yang penting dari hasil *tokenizing*. Setelah itu dilakukan proses *stemming* untuk pengecekan kata dasar yang terdapat pada kamus Bahasa Makassar. Jika kata yang diproses tersebut merupakan kata dasar. Maka tahapan ini akan berlanjut untuk memproses kata berikutnya yang akan dilakukan pengecekan. Jika bukan kata dasar maka proses akan berlanjut pada proses normalisasi.

Kemudian kata akan dilakukan pengecekan dengan menggunakan metode *Damerau Distance*. Metode ini melakukan pengecekan dengan melakukan perhitungan jarak pada *string*. Jika hanya satu kata dengan jarak terpendek maka proses metode ini akan selesai.

3.5 Perancangan

Perancangan digunakan untuk memudahkan dalam pembuatan aplikasi di mana aplikasi ini dirancang berdasarkan hasil analisa yang telah dilakukan pada proses sebelumnya. Adapun tahapan dalam merancang aplikasi ini yaitu:

3.5.1 Flowchart

Pada tahapan ini perancangan dengan menggunakan *flowchart* digunakan untuk menggambarkan bagaimana alur dari sebuah proses algoritma dibuat.

3.5.2 Pseudocode

Pada tahapan ini berisi langkah-langkah algoritma normalisasi pada bahasa Makassar. Tahapan langkah-langkah normalisasi ditulis dalam bentuk *pseudocode*.

3.5.3 Perancangan Database

Database merupakan salah satu komponen penting dalam pembuatan aplikasi karena berisi tabel-tabel data, *field* maupun atribut.

3.5.4 Perancangan Interface

Perancangan *interface* atau antarmuka bertujuan untuk menciptakan tampilan desain aplikasi yang akan dibuat agar pengguna aplikasi dapat dengan mudah memahaminya serta terjalin hubungan yang baik antara pengguna dengan aplikasi.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.6 Implementasi

Pada tahap implementasi ini dilakukan setelah tahap analisa dan perancangan selesai dilakukan, menjelaskan secara rinci bagaimana penerapan sistem yang telah selesai di analisa dan dirancang sebelumnya. Tahapan ini membutuhkan perangkat lunak (*software*) dan perangkat keras (*hardware*):

Hardware yang akan digunakan yaitu:

1. *Processor* : Intel(R) Core(TM) i3
2. *RAM* : 4 GB
3. *HDD* : 500 GB

Software yang akan digunakan yaitu:

1. *Sistem Operasi* : Windows 10
2. *Bahasa Pemrograman* : PHP
2. *Platform* : Microsoft
3. *Browser* : Google Chrome

3.7 Pengujian

Tahapan ini akan dilakukan pengujian terlebih dahulu pada aplikasi yang telah dibangun. Proses pengujian dilakukan untuk melihat apakah aplikasi sudah sesuai atau belum dengan proses implementasi yang dilakukan sebelumnya. Metode pengujian kode program yang digunakan yaitu *White Box* dan pengujian aplikasi dilakukan menggunakan pengujian kata uji dan akurasi. *Validator* adalah seseorang yang menjadi sumber dari pembuatan aplikasi ini. Sumber pada penelitian ini yaitu orang Makassar yang mengerti ataupun yang dapat menggunakan dialek bahasa Makassar.

3.8 Kesimpulan dan Saran

Pada tahap terakhir ini akan diperoleh sebuah kesimpulan dari berdasarkan hasil penelitian yang telah dilakukan. Hasil ini diambil dari tahapan awal hingga ke tahapan akhir yang bertujuan untuk mengetahui penelitian ini sudah dilakukan sesuai dengan apa yang diinginkan sehingga bermanfaat bagi yang akan menggunakannya.



BAB 5 PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian normalisasi teks bahasa Makassar yang telah dilakukan, maka diperoleh kesimpulan sebagai berikut:

1. Implementasi algoritma *Damerau Leveinshtein* ini dapat menutupi kekurangan dari algoritma *Leveinshtein Distance*, karena memiliki 1 tambahan proses edit yaitu proses *transposition* diantara dua karakter dan dapat mengoreksi kesalahan penulisan kata dengan tepat.
2. Hasil pengujian dengan menggunakan rumus *suggestion adequacy* yang didapatkan pada penelitian ini memiliki rata-rata sebesar 72,22%

5.2 Saran

Adapun saran yang dapat digunakan untuk pengembangan selanjutnya adalah dengan:

1. Melengkapi kamus bahasa Makassar- Indonesia.
2. Sebaiknya gunakan data yang sesuai dengan aturan morfologi bahasa Makasar.
3. Sebaiknya perlu menambahkan kamus entitas di dalam *database* yang berisi kata nama orang, nama tempat, unit objek dan lain sebagainya untuk digunakan saat normalisasi atau tidak menemukan kata yang benar.

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Azizurahman, S., Firdaus, Y., & Suryani, A. A. (2011). Analisis Dan Implementasi Metode N-Gram Pada Information Retrieval.
- Bantoro, G. A., Adji, T. B., & Purnamasari, A. E. (2014). Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation. *Citee*, June, 39–43.
- CNN Indonesia. (2018). Badan Bahasa Catat Kini Ada 668 Bahasa Daerah di Indonesia. CNN Indonesia.
- Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171–176. <https://doi.org/10.1145/363958.363994>
- Darwis, P. H. M. (2011). Nasib Bahasa Daerah Di Era Globalisasi.
- Ethnologue. (2019). Languages of Indonesia. Ethnologue.Com.
- Fahma, A. I. (2018). Identifikasi Kesalahan Penulisan Kata (Typographical Error) pada Dokumen Berbahasa Indonesia Menggunakan Metode N-gram dan Levenshtein Distance. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(1), 53–62.
- Hestiani, D. R. (2012). Semangat, Verba Berafiks Bahasa Jawa dalam Rubrik Cerita Rakyat “Pasir Luhur Cinatur” pada Majalah Panjebar. S1 Thesis, Universitas Negeri Yogyakarta.
- Irwansyah, N., Mulyani, S., & Rokhayati, R. (2017). The Use of Batavia language in South Jakarta (Study of sociodialectology). 1(2), 126–130.
- Kementerian Pendidikan dan Kebudayaan Republik Indonesia. (2016). Pengertian Bahasa Menurut Istilah.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lisangan, E. A. (2015). Implementasi n-gram Technique dalam Deteksi Plagiarisme pada Tugas IMPLEMENTASI n-GRAM TECHNIQUE DALAM DETEKSI. May.

Lisangan, E. A. (2013). Natural Language Processing Dalam Memperoleh Informasi Akademik Mahasiswa Universitas Atma Jaya Makassar. Jurnal Tematika.

Fusetti, M., Ruzsics, T., Göhring, A., Samardžić, T., & Stark, E. (2018). Encoder-Decoder Methods for Text Normalization. Proceedings of the Fifth Workshop on {NLP} for Similar Languages, Varieties and Dialects ({V}ar{D}ial 2018), 1998, 18–28. <https://www.aclweb.org/anthology/W18-3902>

Maghira, T., Cholissodin, I., & Widodo, A. (2017). Deteksi Kesalahan Ejaan dan Penentuan Rekomendasi Koreksi Kata yang Tepat Pada Dokumen Jurnal JTIK Menggunakan Dictionary Lookup dan Damerau-Levenshtein Distance. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 1(6), 498–506.

Manyambeang, A. K., Mulya, A. K., & Nasruddin. (1996). Tata bahasa Makassar. Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan.

Maryam, S. (2016). Analisis Morfologi Bentuk Pasif Bahasa Jawa Banyumas. Jurnal Unmuh Jember.

Ramadhanti, F., Wibisono, Y., & Sukanto, R. A. (2019). Analisis Morfologi untuk Menangani Out-of-Vocabulary Words pada Part-of-Speech Tagger Bahasa Indonesia Menggunakan Hidden Markov Model. Jurnal Linguistik Komputasional (JLK), Vol. 2, No. 1, Maret 2019.

Rosyid, M. (2013). Punahnya Bahasa Di Tengah Usaha Mengeksiskan Aksara Nusantara.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Sari, Y. A., & Adikara, P. P. (2019). Klasifikasi Hate Speech Berbahasa Indonesia di Twitter Menggunakan Naive Bayes dan Seleksi Fitur Information Gain dengan Normalisasi Kata. 3(5), 4914–4922.

Manjuntak, M. S., Sujaini, H., & Safriadi, N. (2018). Spelling Corrector Bahasa Indonesia dengan Kombinasi Metode Peter Norvig dan N-Gram. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(1), 17. <https://doi.org/10.26418/jp.v4i1.24075>

Sutisna, U., & Adisantoso, J. (2010). Koreksi Ejaan Query Bahasa Indonesia Menggunakan Algoritme Damerau Levenshtein. *Jurnal Ilmiah Ilmu Komputer*, 8(2), 245772.

Yulianto, M. M., Arifudin, R., & Alamsyah, A. (2018). Autocomplete and Spell Checking Levenshtein Distance Algorithm To Getting Text Suggest Error Data Searching In Library. *Scientific Journal of Informatics*, 5(1), 75. <https://doi.org/10.15294/sji.v5i1.14148>

LAMPIRAN A

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Syaifullah
 Alamat : Jln. Halimun No.17, Bukit Baruga, Makassar - Sulawesi Selatan
 Pekerjaan : Karyawan swasta
 Bahasa Sehari-hari : Indonesia dan Makassar

Dengan ini menyatakan bahwa saya bersedia menjadi Validator pada penelitian:

Nama : Mujiburahman
 Nim : 11451104625
 Jurusan : Teknik Informatika
 Judul Penelitian : Normalisasi Teks Bahasa Makassar Menggunakan Damerau Levenshtein

Demikian Surat Pernyataan ini saya buat untuk digunakan seperlunya, atas perhatian saya ucapkan terimakasih.

Makassar, September 2021


 (..... Syaifullah)

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN B

Berikut merupakan tabel perbandingan hasil mesin dan hasil pakar.

No	Kata Tidak Baku	Hasil Mesin	Hasil Validator	Nilai
1	palla	balla, jalla, kalla, palla, pallu	balla	1
2	dg	di	di	1
3	suarana	suarang	Suara'	-0,5
4	lontara	lontarak	lontarak	1
5	hajji	hakji	hakji	1
6	oh	o, oe	oe	0,5
7	karaeng	karaeng, maraeng	karaeng	1
8	segang	selang	selang	1
9	ini	ani, inti	ani	1
10	bhs	bas	basa	-0,5
11	rammusuki	rammusuk	rammusuk	1
12	ajji	aji	aji	1
13	tabe	jabe, taba, tabek	tabek	0,5
14	dmna	dana	dana	1
15	sya	sa	sa	1
16	suara	siara, suarak, sura	sura	0,5
17	nakku	nakke, nakkuk	nakkuk	0,5
18	punya	punna	punna	1
19	agne	angge, anne	anne	0,5
20	tul	tuli, tulu, tuo	tuli	1
21	lalanga	lalang	lalang	1
22	matea	mate, matua	mate	1
23	nakkuku	nakkuk	nakkuk	1
24	khi	koi	koi	1
25	dulu	dudu, tulu, ulu	ulu	0,5
26	sinrili	sinrilik	sinrilik	1
27	ada	adak, ala, ama, apa, ara, dada, ida, lada, nada, pada, wada	adak	1
28	peduli	paduli	paduli	1
29	ya	ia, ja, la, na, sa	ia	1
30	ka	ia, ja, la, na, sa	la	0,5
31	salamaki	salamak	salamak	1
32	lino	libo, lio, sino	sino	0,5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

33	aminn	aming	aming	1
34	orang	erang, irang, onrang, opang, urang	urang	0,5
35	pangkep	pangke	pangke	1
36	ku	au, uk	uk	0,5
37	jhi	jai, ji	ji	0,5
38	itu	ibu	ibu	1
39	mama	ama, gama, hama, jama, macca, maka, maro, mamo, pama, rama, sama, tama	ama	1
40	tolong	bolong, kolong, polong, lolong, solong,	solong	0,5
41	toli	loli, tali, toali, tolo, tuli	tuli	0,5
42	jaman	jama, jamak	jamang	-0,5
43	mauka	maka	maka	1
44	ciba	cika, cina, coba, liba, riba	coba	0,5
45	ke	ce, e, oe	e	0,5
46	mu	au	au	1
47	ello	allo, elo	elo	0,5
48	ki	aki, di, I, ji, ki, koi, ri, si	si	0,5
49	nek	bek, dek, iek, lek	dek	0,5
50	tpi	tai, tapi	tapi	0,5
51	skrang	suarang	suarang	1
52	mantan	mantang, mantar	mantang	1
53	pele	kale, lele, pake, palu, pare, pase, pile	pake	0,5
54	pak	cak, gak, hak, jak, paik, pake, paku, pau, pauk, pik, sak, tak, uak, upak	paku	0,5
55	nya	na	na	1
56	makan	maka	maka	1
57	ca	cak, cao, cau, cek, coa, ia, ja, la, na, sa	cak	1
58	nata	bata, jaba, naba, nada, naga, naia, nana, nawa, pata, rata	bata	1
59	ggala	gala	gala	1
60	tamanga	tamang	tamang	1
61	dg	di	di	1
62	benang	berang, bonang, penanag, tenang	berang	1
63	jalan	jala, jalak, jalang	jalang	0,5
64	manaki	manak	manak	1
65	memangg	memang	memang	1
66	salam	sala, salak	sala	1
67	dari	ari, barri, dara, jari, kari, lari, mari	jari	0,5
68	kereng	pereng	pereng	1
69	burasa	burasak	burasak	1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



70	mantong	mantang, pantong	mantang	1
71	taww	tawa	tawa	1
72	tambah	tamba, tambak	tamba	1
73	biar	bira	bira	1
74	tongko	longko, tongka, tongkok	tongkok	0,5
75	mi	di, I, ji, ki, ri, si	ki	0,5
76	ta	ia, ja, la, na, sa, tai, tak, tau, tea, toa	tau	0,5
77	baruka	barua, baruga	baruga	0,5
78	ka	ia, ja, la, na, sa	sa	0,5
79	ke	ce, e, oe	e	0,5
80	cocoki	cocok	cocok	1
81	sekali	sakali	sakali	1
82	doe	boe, doek, oe, roe	doek	0,5
83	untung	kuntung, runtung, unjung, utung	utung	0,5
84	pi	di, I, ji, ki, pia, pik, ri, sa	pik	0,5
85	sede	dede, sese	sedek	-0,5
86	colek	colek, lolek	colek	1
87	galla	balla, gala, jalla, kalla	gala	0,5
88	nonci	konci	konci	1
89	lamami	lamari	lamari	1
90	lgi	lagi	lagi	1
91	coto	coko, codo, colo, conto, coro, oto, poto	coro	0,5
92	andalan	andanlang	andanlang	1
93	itu	ibu	iruk	-0,5
94	abo	ao, aro, awo, ayo, labo	labo	0,5
95	ada	adak, ala, ama, apa, ara, dada, ida, lada, nada, pada, wada	adak	1
96	ma	ama, ia, ja, la, mae, mea, na, sa	mae	0,5
97	gue	gae, gua, kue, pue, sue	gua	0,5
98	mami	mamo, mari	mari	0,5
99	tongk	aong, atong, bong, gong, tang, toeng, tongi	tang	0,5
100	sagang	agang, bagang, sabang, sadang, sahang, sakgang, salang, samang, sayang, siagang	siagang	0,5
101	jangan	jangang	jangang	1
102	kadong	dodong, kalong, kodong	kodong	0,5
103	kah	wah	wah	1
104	teak	geak, tak, tea, tepak, terak, tuak	tak	0,5
105	belah	bela, belak	bela	1
106	biak	bidak, bisak, buak, niak	buak	0,5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

107	papua	papa	papa	1
108	beee	bebe, bene, bese	bebe	1
109	main	maik	maik	1
110	dana	dana, dang	dana	1
111	kak	akak, cak, ekak, gak, hak, jak, kako, koak, sak, tak, uak,	tak	0,5
112	gamasi	galasi	galasi	1
113	jg	ja, ji	ji	0,5
114	bro	aro, biro	biro	0,5
115	mantap	mantar	mantar	1
116	aparat	aparak	aparak	1
117	kak	akak, cak, ekak, gak, hak, jak, kako, koak, sak, tak, uak,	tak	0,5
118	laa	ala, la, laba, lada, laga, laja, lara, lasa, lau, lawa, lao	la	1
119	de	cek, dek, di, e, oe	dek	0,5
120	kanda	canda, danda, ganda, panda	ganda	0,5
121	dan	dana, dang	dana	1
122	jie	ji, jik	ji	1
123	bamma	bamba, kamma, lamma, ramma	kamma	0,5
124	maki	aki, baki, lari, maik, maka, mari, raki	mari	0,5
125	ta	ia, ja, la, na, sa, tai, tak, tau, tea, toa	tak	0,5
126	haha	hama, hawa	hawa	0,5
127	ih	I, ia, io	i	1
128	eh	e	e	1
129	jadi	jai, jari, jati	jari	0,5
130	urangi	ukrangi, urang	ukrangi	1
131	pelle	pele	pelek	-0,5
132	yoko	boko, koko, koko, loko, soko, toko	boko	1
133	be	bea, bek, beo, ce, e, oe	bek	0,5
134	ol	o, oe, olo	olo	0,5
135	haaa	hama, hawa	hawa	0,5
136	kena	pena, tena	kenna	-0,5
137	dua	duda, dupa, duta, gua, kua, pua, rua	duda	1
138	ko	ao, io, koi, o	koi	0,5
139	bapa	apa, baca, bala, bara, bata, baya, papa, rapa, sapa, tapa	bala	0,5
140	tawwa	tawa	tawa	1
141	we	ce, e, oe	e	0,5
142	to	ao, io, ko, oto, toa, tuo	tuo	0,5
143	nu	anu, au, na	anu	1

144	ciniki	cinik	cinik	1
145	masa	lasa, maca, maka, mara, masak, mase, maso, musa, rasa, sasa	masak	0,5
146	kah	wah	wah	1
147	poko	boko, koko, koko, loko, poka, poka, pokok, polo, ponko, poro, poso, poto, soko, toko	boko	1
148	sy	sa, si	si	0,5
149	nya	na	na	1
150	pakanak	pakanag	pakkanang	-0,5
151	sdeng	seng	seng	1
152	ma	ama, ia, ja, la, mae, mea, na, sa	mae	0,5
153	cepat	cepat	cepat	1
154	sappo	cappo, lappo, rappo, sampo, sappa, sappe	cappo	1
155	nanas	nana	nana	1
156	adama	agama	agama	1
157	peccei	pecce, pacci	pacce	-0,5
158	dottoro	dottorok	dottorok	1
159	nia	ia, na, nai, naia, niak, nila, nipa, pia	ia	1
160	masseng	massang, massing	massang	1
161	sekke	sekek, sekkek	sekkek	1
162	siapa	sapa, siara, sikapa	sapa	1
163	deng	daeng, dang, reng, seng	daeng	1
164	mau	au, maucau, kau, lau, mae, maju, maru, pau, rau, sau, tau	kau	0,5
165	lawan	lawa	lawa	1
166	tojeng	toeng, tokeng	tokeng	0,5
167	eee	ece, ere	ere	0,5
168	beee	bea, bebe, bek, bene, bese, boe	bebe	0,5
169	nda	ida, na, nada	nada	0,5
170	pale	kale, lele, pake, palu, pare, pase, pele, pile aji, bai, bajik, baju, baki, bakji, bali, bani, bari, bari, bati, gaji, taji	palu	0,5
171	baji		baju	0,5
172	aren	are, areng	areng	0,5
173	agak	abak, adak, ahak, akak, laka, anak, arak, asak, atak, ayak, gagak, gak, lagak	akak	0,5
174	ballasa	ballasak	ballasak	1
175	deh	dek	dek	1
176	bawa	bacca, bala, bara, bata, baya, hawa, kama, lawa, nawa, rawa, tawa	bara	0,5
177	ganaja	ganja	ganja	1
178	puang	buang, pung, suang, taung, uang	pung	0,5

- Hak Cipta Dilindungi Undang-Undang**
- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

179	pakana	pakanag	pakana-kana	-0,5
180	tanta	banta, tahta, tanpa, tantu	tantu	0,5
181	anni	ani, anne, anti, manni	manni	0,5
182	tongji	tongi	tongi	1
183	dii	didi	didi	1
184	sde	sue	sue, sede	1
185	ammale	ammane	ammalli	-0,5
186	allea	alle	alle	1
187	twwa	tawa	tawwa	-0,5
188	jki	aki, jai, ji, jik	jaki	-0,5
189	panraki	panrak	panrak	1
190	lambung	lambang, lamung, limbung, sambung, tambung	lambang	1
191	ber	bea, bek,beru	beru	0,5
192	sembarang	sambarang	sambarang	1
193	sangat	sangak	sangak	1
194	odo	bodo, sudo, odok, olo, oto, sodo	bodo	1
195	tdk	tak	tak	1
196	dl	di	di	1
197	dr	di	di	1
198	sede	sese, sase	sese	1
199	mannang	annang, bannang, jannang, manang, manngang, mantang, pannang, sannang	manang	0,5
200	bos	bas, boe, bos, los	bosi	0,5
201	dosen	doseng	doseng	1
202	jt	ji, jak	ji	1
203	ayanga	ayang	ayang	1
204	jarrako	jarrak	jarrak	1
205	mauko	muko	muko	1
206	neneka	nenek	nenek	1
207	janganga	jangang	jangang	1
208	ipi	ipik, pipi	pipi	0,5
209	jaina	janna, jina	janna	1
210	sukski	sukki	sukki	1
211	sepe	lepe, pepe, sape, sese	sape	0,5
212	gere	ere, jere, kere, pere, rere	rere	0,5
213	tompi	tumpi	tumpi	1
214	tante	rante, rantu	rante	1
215	iya	ia, ida, ija, isa, riya	ia	1
216	yoko	boko, coko, koko, loko, soko, toko	boko	1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

217	bapa	apa, baca, bala, bara, bata, baya, papa, rapa, sapa, tapa	apa	1
218	poko	boko, koko, koko, loko, poka, poke, pokok, polo, ponko, poro, poso, poto, soko, toko	boko	1
219	tua	gua, kua, pua, rua, tau, tea, toa, tuak, tuda, tuna, tuo	gua	1
220	nah	na, nai, wah	na	1
221	macam	maca	maca	1
222	ca	cak, coa, cau, ce, coa, ia, ja, la, na, sa	cau	0,5
223	bajika	bajik	bajik	1
224	katte	batte	batte	1
225	baugang	bagang	bagang	1
226	tonga	conga, jonga, toga, tongi, tongka	tongka	0,5
227	baek	balek, barek, bauk, bek,	balek	1
228	sull	sule, sulle, sulo	sulle	0,5
229	ok	iok, jok, o, oe, uk	jok	0,5
230	porei	pore	pore	1
231	mentong	mengong, menteng	menteng	0,5
232	an	ane, ani, anu, ao, au, na	na	0,5
233	skak	akak, ekak, sak, sikak, suak, suak	akak	1
234	mat	mae, mate, mato	mate	0,5
235	andalanga	andanlang	andalang	-0,5
236	inu	anu, ibu	anu	1
237	makkala	takkala	makkalak	-0,5
238	ong	aong, bong, gong	aong	1
239	gonrong	lonrong, ronrong	ronrong	1
240	toloi	tolo, tolok	tolo	1
241	abo	ao, aro, awo, ayo, labo	labo	0,5
242	adama	agama	agama	1
243	peccei	pacce, pacci	pacce	1
244	moke	mode, mole, poke, toke	mole	0,5
245	akkala	takkala	takkala	1
246	klo	elo, olo	olo	0,5
247	ke	ce, e, oe	e	0,5
248	mami	mamo, mari	mari	0,5
249	jangan	jangang	jangang	1
250	lurang	kurang, larang, lerang, urang, yurang	kurang	1
251	bangkok	bangko, bengkok	bengkok	0,5
252	misseng	isseng	isseng	1
253	kapang	apang, dapang, japang, rapang	apang	1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

254	jolie	jole	jole	1
255	jih	ji, jik	ji	1
256	andalanga	andanlang	andalang	-0,5
257	dapa	apa, dada, dana, dara, dupa, papa, rapa, sapa, tapa	rapa	0,5
258	suka	buka, saka, suak, sukak, suku, sura, susa	buka	1
259	lucu	cucu	cucu	1
260	kocak	koak, Kocak, pocak, rocak	koak	1
261	jga	ja, jaga	jaga	0,5
262	nn	na	na	1
263	raja	gaja, jaja, laja, paja, raga, rama, rapa, rasa, rata, rawa, raya, taja	taja	0,5
264	jangko	bangko, jangka, jangki	jangki	0,5
265	injak	anjak, ijak, ingak, inja, insak, tinjak	tinjak	0,5
266	om	oe, o	o	0,5
267	palukka	palukkak	palukkak	1
268	toh	toa	toa	1
269	dompala	dompalak	dompalak	1
270	ni	ani, di, I, ji, ki, na, nai, ri, si	nai	0,5
271	bodoi	bodo	bodo	1
272	de	ce, dek, di, e, oe	dek	0,5
273	ada	adak, ala, ama, apa, ara, dada, ida, lada, nada, pada, wada	adak	1
274	lantang	antang, bantang, gantang, landang, lattang, lontang, mantang, mantang, rantang, tantang	lantang	0,5
275	bangngi	banngi	bangngi	-0,5
276	boss	bosi	bosi	1
277	sagang	agang, bagang, sabang, sadang, sahang, sakgang, salang, samang, sayang, siagang	sakgang	0,5
278	kita	gita, pita, rita, sita	pita	0,5
279	toongi	tongi	tongi	1
280	porei	pore	pore	1
281	batena	bateng	bate	0,5
282	meki	meko	meko	1
283	bayu	babu, baju, balu, baru, batu, bau, baya, kayu, sayu	baju	0,5
284	nanti	anti, ganti	anti	1
285	logat	logak	logak	1
286	nda	ida, na, nada	nada	0,5
287	iya	ia, ida, ija, isa, riya	ia	1
288	asseng	aseng, assang, assung, isseng	aseng	1



289	salla	balla, jalla, kalla, sala, sallak, sallo	sallak	0,5
290	lipaki	lipak	lipak	1
291	dudui	dudu	dudu	1
292	ipi	ipik, pipi	ipik, pipi	1
293	mako	kako, mado, maka, malo, mamo, mato, meko, muko, sako	mado	0,5
294	allei	alle	alle	1
295	foto	oto, poto	poto	0,5
296	bareng	areng, baeng, barang, bateng, lareng,	baeng	0,5
297	palla	balla, jalla, kalla, pallak, pallu	pallak	0,5

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR RIWAYAT HIDUP

Informasi Personal

	Nama	: Mujiburahman
	Tempat Lahir	: Bekawan
	Tanggal Lahir	: 30 Juni 1996
	Agama	: Islam
	Jumlah Bersaudara	: 6
	Anak ke	: 6
	Kebangsaan	: WNI

Alamat

Alamat	: Jl. Suka Karya, Kec. Tampan, Kota Pekanbaru, Riau
E-mail	: mujiburahman@students.uin-suska.ac.id

Riwayat Pendidikan

Tahun 2002 - 2008	: MIS Darussalam Bekawan
Tahun 2008 - 2011	: MTs Bekawan
Tahun 2011 - 2014	: SMKN 2 Tembikahan
Tahun 2014 - 2021	: Teknik Informatika UIN Suska Riau

Informasi Orang Tua

Ayah	: Denni
Pendidikan Terakhir	: Sekolah Dasar
Ibu	: Siti Hajerah
Pendidikan Terakhir	: Sekolah Dasar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.