

**ANALISIS PERBANDINGAN ALGORITMA NAZIEF &
ADRIANI DENGAN ALGORITMA ARIFIN SETIONO UNTUK
STEMMING BAHASA INDONESIA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

HILMI AZIIZI

16.11.0557

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

**ANALISIS PERBANDINGAN ALGORITMA NAZIEF &
ADRIANI DENGAN ALGORITMA ARIFIN SETIONO UNTUK
STEMMING BAHASA INDONESIA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

HILMI AZIIZI

16.11.0557

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PERBANDINGAN ALGORITMA STEMMING NAZIEF &
ADRIANI DENGAN ALGORITMA STEMMING ARIFIN SETIONO
UNTUK STEMMING BAHASA INDONESIA**

yang disusun dan diajukan oleh

Hilmi Aziizi

16.11.0557

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 31 Juli 2023

Dosen Pembimbing,



Windha Mega Pradnya D. M. Kom

NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS PERBANDINGAN ALGORITMA STEMMING NAZIEF &
ADRIANI DENGAN ALGORITMA STEMMING ARIFIN SETIONO
UNTUK STEMMING BAHASA INDONESIA**

yang disusun dan diajukan oleh

Hilmi Aziizi

16.11.0557

Telah dipertahankan di depan Dewan Penguji
pada tanggal 31 Juli 2023

Susunan Dewan Penguji

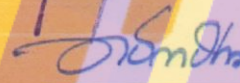
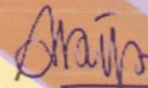
Nama Penguji

Arif Dwi Laksito, M.Kom
NIK. 190302150

Uyock Anggoro Saputro, M.Kom
NIK. 190302419

Windha Mega Pradnya D, M.Kom
NIK. 190302185

Tanda Tangan



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 31 Juli 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hilmi Aziizi
NIM : 16.11.0557

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Perbandingan Algoritma Nazief & Adriani Dengan Algoritma Arifin Setiono Untuk Stemming Bahasa Indonesia

Dosen Pembimbing : Windha Mega Pradnya D, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 31 Juli 2023

Yang Menyatakan,



Hilmi Aziizi

HALAMAN PERSEMBAHAN

Puji syukur kehadirat Allah SWT, atas limpahan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan skripsi ini. Skripsi ini saya persembahkan untuk:

1. Kedua orang tua saya yang selalu memberikan bantuan baik dari segi bantuan materi ataupun bantuan moral. Berkat doa mereka skripsi ini dapat terselesaikan.
2. Ibu Windha Mega Pradnya D, M.Kom, selaku dosen pembimbing yang telah memberikan kesempatan, harapan, dan telah sabar dalam membimbing saya selama proses pengerjaan skripsi ini.
3. Seluruh dosen Universitas Amikom Yogyakarta yang telah memberikan banyak sekali ilmu serta pelajaran yang sangat membantu dalam proses penyusunan skripsi ini.
4. Teman-teman dari S1 IF-09 Universitas Amikom Yogyakarta yang telah mendukung saya dari awal sampai akhir.

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Analisis Perbandingan Algoritma Nazief & Adriani Dengan Algoritma Arifin Setiono Untuk Stemming Bahasa Indonesia. Shalawat serta salam senantiasa tercurah kepada junjungan kita Nabi Muhammad SAW, yang telah memberikan teladan yang luar biasa dalam segala aspek kehidupan.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata 1 dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini, Maka pada kesempatan ini saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Windha Mega Pradnya D, M.Kom. selaku Dosen Pembimbing yang telah memberikan pengarahan, masukan, dan dorongan yang sangat membantu dalam proses pembuatan skripsi ini.

Yogyakarta, 31 Agustus 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PENYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur.....	7
2.2 Dasar Teori.....	11
2.2.1 Text Mining.....	11
2.2.2 Text Preprocessing.....	11
2.2.3 Case Folding.....	11
2.2.4 Tokenizing.....	11
2.2.5 Stop Word Filtering.....	12
2.2.6 Stemming.....	12
2.2.6.1 Algoritma Nazief-Adriani.....	13

2.2.6.2 Algoritma Arifin Setiono.....	19
2.2.7 Aplikasi Berbasis Web	20
2.2.7.1 Pengenalan PHP	21
2.3 Metode Pengujian.....	21
2.3.1 Word Conflation Class	21
2.3.2 Akurasi	22
2.3.3 Waktu Proses	22
BAB III METODE PENELITIAN.....	22
3.1 Prosedur Penelitian.....	23
3.2 Data Penelitian	24
3.2.1 Persiapan Data.....	24
3.2.2 Preprocessing Text	24
3.2.2.1 Stop Word Filtering.....	24
3.2.2.2 Case Folding	24
3.2.2.3 Tokenizing.....	24
3.2.2.4 Penerapan Algoritma Stemming Nazief&Adriani	25
3.2.2.5 Penerapan Algoritma Stemming Arifin Setiono.....	26
3.3 Alat/Instrumen.....	27
3.3.1 Kebutuhan Perangkat Keras.....	27
3.3.2 Kebutuhan Perangkat Lunak.....	27
3.4 Metode Analisis.....	28
3.4.1 Analisis Masalah	28
3.4.2 Analisa Fungsional.....	28
3.4.3 Analisa Non-Fungsional.....	28
3.5 Penjabaran Perancangan.....	29
3.5.1 Perancangan Proses.....	29
BAB IV IMPLEMENTASI DAN PEMBAHASAN.....	32
4.1 Persiapan Data.....	32

4.2 Preprocessing Text	33
4.2.1 Stop Word Filtering	33
4.2.2 Case Folding	35
4.2.3 Tokenizing	37
4.3 Penerapan Algoritma	38
4.3.1 Algoritma Nazief&Adriani	38
4.3.2 Algoritma Arifin Setiono	41
4.4 Parameter Analisi Hasil	43
4.4.1 Word Conflation Class	43
4.4.2 Akurasi	44
4.4.3 Akurasi Waktu Proses	45
BAB V KESIMPULAN DAN SARAN	50
5.1 KESIMPULAN	50
5.2 SARAN	51
DAFTAR PUSTAKA	52

DAFTAR TABEL

Tabel 2.1 Table Perbandingan Penelitian	7
Tabel 2.2 Imbuhan Yang Tidak Diizinkan.....	14
Tabel 2.3 Aturan Penghapusan Awalan Me-	15
Tabel 2.4 Aturan Penghapusan Awalan Pe-.....	16
Tabel 2.5 Aturan Penghapusan Awalan Be-.....	16
Tabel 2.6 Aturan Penghapusan Awalan Te-.....	17
Tabel 4.1 Stop Word Filtering	32
Tabel 4.2 Case Folding.....	34
Tabel 4.3 Tokenizing	36
Tabel 4.4 Hasil Stemming Algoritma Nazief&Adriani.....	38
Tabel 4.5 Hasil Stemming Algoritma Arifin Setiono	40
Tabel 4.6 Akurasi Kedua Algoritma.....	43
Tabel 4.7 Akurasi Waktu	44
Tabel 4.8 Hasil Stemming Semua Data.....	47

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	22
Gambar 3.2 Flowchart Algoritma Stemming Nazief&Adriani	24
Gambar 3.3 Flowchart Algoritma Stemming Arifin Setiono	25
Gambar 3.4 Flowchart Keseluruhan Sistem.....	28
Gambar 3.5 Flowchart Preprocessing Text	29
Gambar 3.6 Flowchart Filtering Stop Word.....	30
Gambar 4.1 Data Text Stemming	31
Gambar 4.2 Implementasi Stemming Algoritma Nazief&Adriani	39
Gambar 4.3 Implementasi Stemming Algoritma Arifin Setiono.....	42
Gambar 4.4 Waktu Proses Algoritma Nazief&Adriani	44
Gambar 4.5 Waktu Proses Algoritma Arifin Setiono	44
Gambar 4.6 Uji Stemming “KalimatUji2.txt” Algoritma Nazief&Adriani	45
Gambar 4.7 Uji Stemming “KalimatUji2.txt” Algoritma Arifin Setiono	46
Gambar 4.8 Uji Stemming “KalimatUji3.txt” Algoritma Nazief&Adriani	46
Gambar 4.9 Uji Stemming “KalimatUji3.txt” Algoritma Arifin Setiono	47

INTISARI

Kebutuhan akan informasi sangat meningkat. Banyak informasi yang perlu diperoleh dengan cepat, dan relevan sesuai kebutuhan, tetapi untuk mendapatkan informasi yang relevan dan akurat sangatlah sulit, dibutuhkan pengolahan kata untuk mendapatkan informasi yang lebih relevan dan akurat. Dalam proses pengolahan kata, terdapat proses stemming, yaitu proses untuk mendapatkan kata dasar dengan menghapus afiks dari kata yang dicari. stemming memainkan peran penting dalam pengolahan kata, membantu pengurangan kata menjadi bentuk akarnya dan meningkatkan efisiensi dan akurasi, dan membantu dalam analisis teks serta pencarian informasi. Dalam konteks bahasa Indonesia, dua algoritma stemming terkemuka telah muncul sebagai pesaing utama: algoritma stemming Nazief & Adriani dan algoritma stemming Arifin Setiono. Dalam penelitian ini menyajikan studi komparatif yang komprehensif dari kedua pendekatan ini untuk menjelaskan kekuatan, kelemahan, dan penerapannya.

Kata kunci : Stemming, Analisis Perbandingan, Algoritma Stemming Nazief & Adriani, Algoritma Stemming Arifin Setiono.

ABSTRACT

The need for information is greatly increased. A lot of information needs to be obtained quickly, and is relevant as needed, but to get relevant and accurate information is very difficult, text processing is needed to get more relevant and accurate information. In the text processing process, there is a stemming process, which is the process of obtaining basic words by removing affixes from the word being searched for. stemming plays an important role in word processing, assisting in reducing words to their root forms and increasing efficiency and accuracy, and assisting in text analysis and information retrieval. In the context of the Indonesian language, two leading stemming algorithms have emerged as major competitors: Nazief & Adriani's stemming algorithm and Arifin Setiono's stemming algorithm. This research presents a comprehensive comparative study of these two approaches to explain their strengths, weaknesses, and applications.

Keyword : *Stemming, Comparative Analysis, Nazief & Adriani Stemming Algorithm, Arifin Setiono Stemming Algorithm.*

