

**IMPLEMENTASI ALGORITMA STEMMING NAZIEF DAN ADRIANI
PADA PREPROCESSING ALGORITMA RABIN KARP UNTUK
STEMMING BAHASA INDONESIA**

SKRIPSI



disusun oleh

Firman Panti Gelen

15.11.8957

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**

**IMPLEMENTASI ALGORITMA STEMMING NAZIEF DAN ADRIANI
PADA PREPROCESSING ALGORITMA RABIN KARP UNTUK
STEMMING BAHASA INDONESIA**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana pada
Program Studi Informatika



**Oleh
Firman Panti Gelen
15.11.8957**

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**

PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA STEMMING NAZIEF DAN ADRIANI
PADA PREPROCESSING ALGORITMA RABIN KARP UNTUK
STEMMING BAHASA INDONESIA**

yang dipersiapkan dan disusun oleh

Firman Panti Gelen
15.11.8957

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Maret 2018

Dosen Pembimbing,



Heri Sismoro, M.Kom
NIK. 190302017

PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA STEMMING NAZIEF DAN ADRIANI
PADA PREPROCESSING ALGORITMA RABIN KARP UNTUK
STEMMING BAHASA INDONESIA**

yang dipersiapkan dan disusun oleh

Firman Panti Gelen

15.11.8957

telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Maret 2019

Susunan Dewan Penguji


Nama Penguji

Tanda Tangan

Anggit Dwi Hartanto, M.Kom
NIK. 190302163

Wiwi Widavani, M.Kom
NIK. 190302272

Heri Sismoro, M.Kom
NIK. 190302017



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Maret 2019

DEKAN FAKULTAS ILMU KOMPUTER



Krisnawati, S.Si, M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institus pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 27 Maret 2019



Firman :

NIM. 15.11.8957

MOTTO

“Hidup ini terlalu singkat untuk terus berdiam dan bermimpi, mulai bertindak, tetap optimis dan mencoba menyelesaikan masalah-masalah serta selalu bersyukur apapun jalan hidup sudah ditentukan oleh Nya”

- Firman Panti Gelen -



PERSEMBAHAN

Pada halaman persembahan ini penulis mempersembahkan dan mengucapkan terima kasih kepada pihak-pihak yang telah membantu dalam penyusunan skripsi ini.

1. Terimakasih dan syukur kepada Allah SWT yang memberikan nikmat dan kasih sayangnya yang tak terhingga sampai sejauh ini.
2. Terimakasih kepada Ibunda tercinta dan Ayah tercinta, yang sudah bekerja keras, selalu memberikan dukungan, doa, serta semangat.
3. Terimakasih kepada dosen pembimbing, Bapak Heri Sismoro, S. Kom, M.Kom yang telah memberikan bimbingannya dalam penyusunan karya tulis ilmiah ini.
4. Terimakasih kepada keluarga besar yang selalu memberikan dukungan dan doa.
5. Terimakasih kepada Asti yang membantu pada saat proses penyusunan skripsi serta sahabat, teman kontrakan, dan teman-teman 15-S1TI-07 yang selalu memberikan bantuan, semangat, dan mengingatkan dalam keadaan apapun.
6. Terimakasih kepada saya sendiri yang telah berjuang dan pantang menyerah untuk mendapatkan apa yang saya inginkan.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakaatuh

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan limpahan rahmat, kemudahan, kelancaran dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan cukup baik. Tidak lupa sholawat serta salam selalu dicurahkan kepada nabi besar dan rasul junjungan kita Nabi Muhammad SAW beserta keluarganya.

Skripsi ini merupakan salah satu bentuk persyaratan kelulusan jenjang Program Strata satu (S1) jurusan Teknik Informatika pada Universitas Amikom Yogyakarta. Penulis menyadari bahwa penulisan skripsi “Implementasi Algoritma Stemming Nazief Dan Adriani Pada Preprocessing Algoritma Rabin Karp Untuk Stemming Bahasa Indonesia” ini masih banyak sekali kekurangan yang itu semua tidak lepas karena keterbatasan penulis.

Dalam pembuatan skripsi ini, tentu saja penulis mendapat banyak bantuan dari berbagai pihak, oleh karena itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Heri Sismoro, S.Kom, M.Kom. selaku Dosen Pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penelitian dan penulisan skripsi ini.
4. Tim penguji, segenap dosen dan karyawan Universitas Amikom yang telah memberikan ilmu dan pengalaman.
5. Kedua orang tua tercinta atas doa, kasih sayang, didikan dan cinta kepada saya.

6. Keluarga besar yang senantiasa mendoakan saya, memberi semangat dan bantuan.
7. Sahabat dan teman kontrakan yang telah memberikan dukungan dan semangat.
8. Teman-teman 15-SITI-07 yang telah menemani selama proses perkuliahan.

Semoga Allah SWT memberikan balasan yang lebih kepada semua yang telah ikut membantu saya dan menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun akan diterima dengan senang hati dan rasa terimakasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Akhir kata penulis ucapkan terima kasih atas kesediaannya untuk membaca dan memahami skripsi ini.

Wasalamu'alaikum Warahmatullahi Wabarakaatuh

Yogyakarta, 2 April 2019

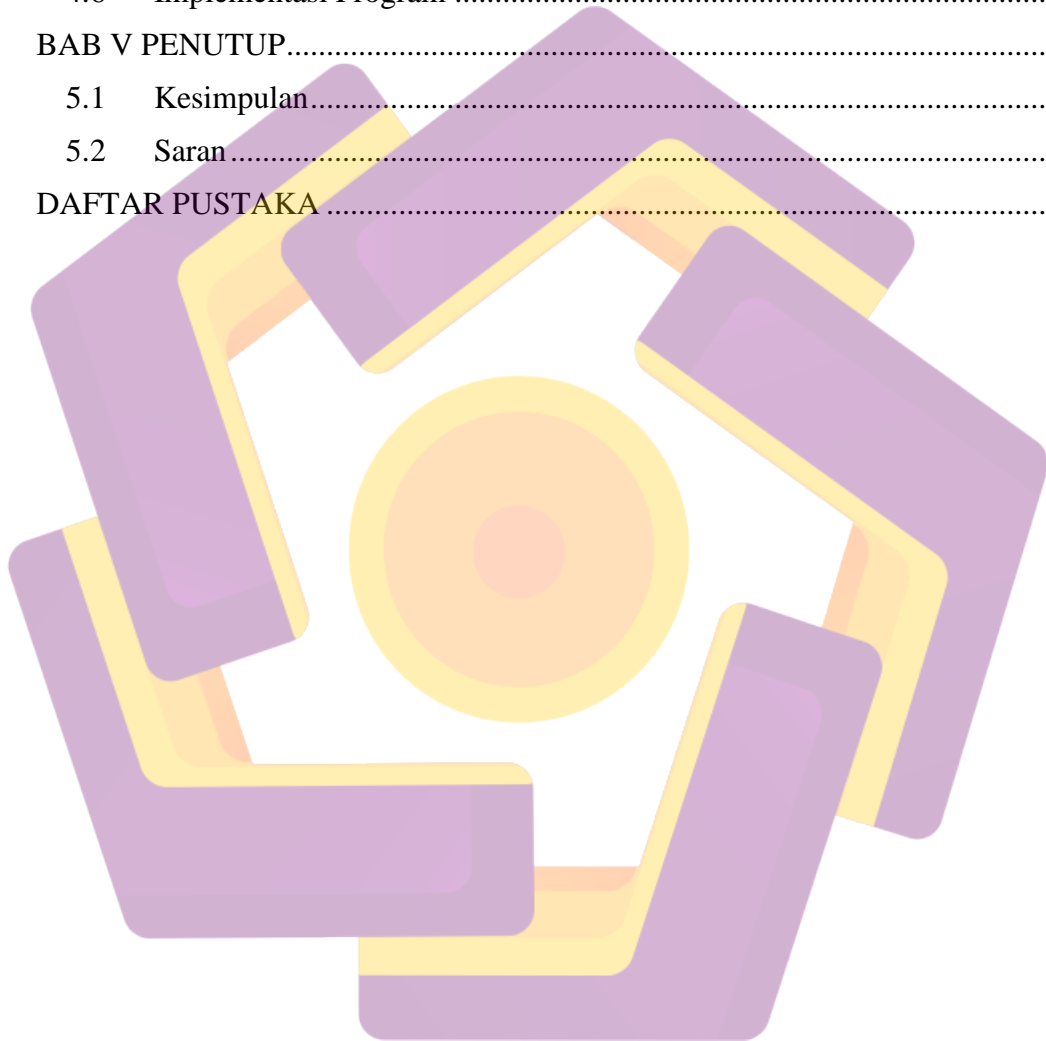
Penulis

DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN.....	iv
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Metode Penelitian.....	4
1.6.1 Metode Pengumpulan data.....	4
1.6.2 Metode Analisis	5
1.6.3 Metode Perancangan	5
1.6.4 Metode Pengembangan	5
1.6.5 Metode Testing.....	5
1.7 Sistematika Penulisan.....	5
1.7.1 BAB I : Pendahuluan	6
1.7.2 BAB II : Landasan Teori.....	6
1.7.3 BAB III : Analisis dan Perancangan	6

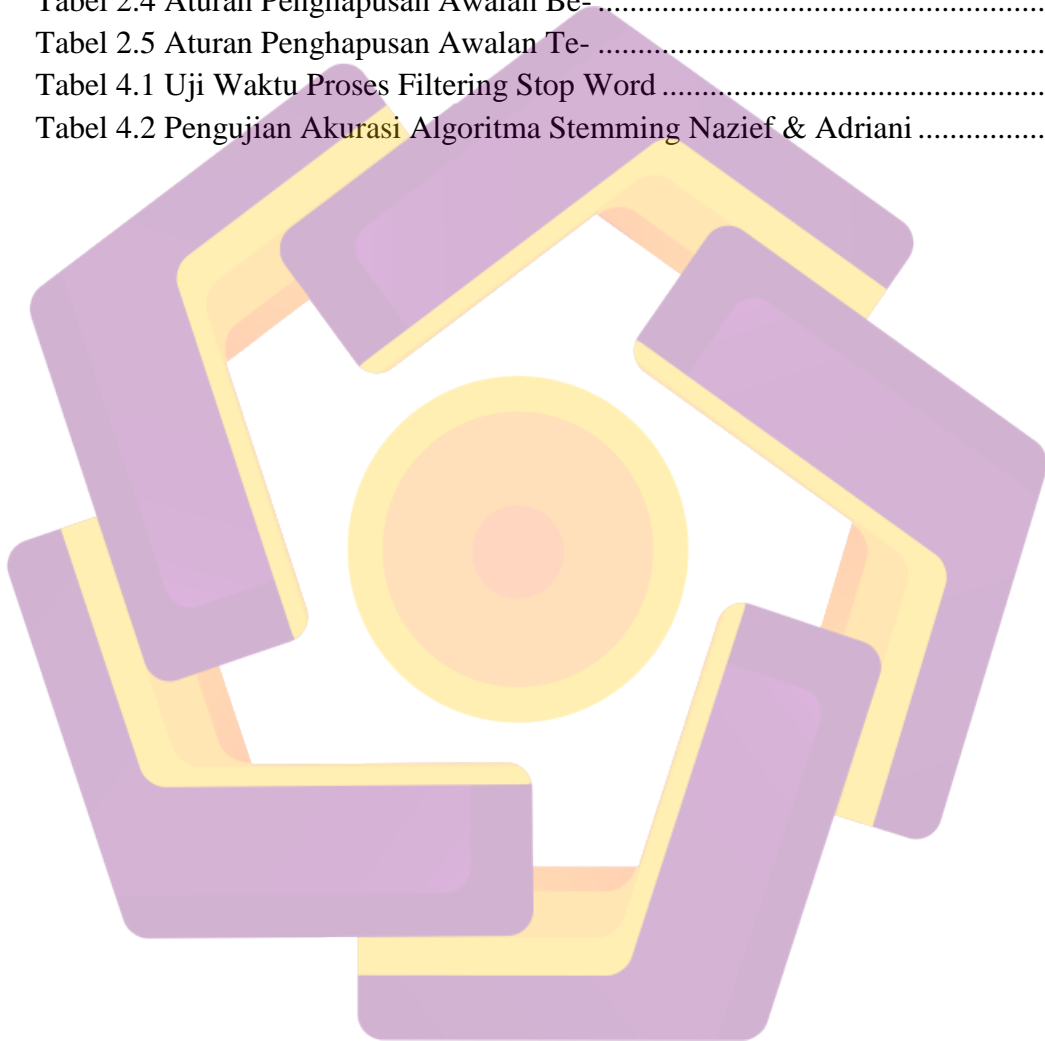
1.7.4	BAB IV : Implementasi dan Pembahasan.....	6
1.7.5	BAB V : Penutup	6
BAB II LANDASAN TEORI		7
2.1	Tinjauan Pustaka	7
2.2	Dasar Teori	10
2.2.1	<i>Text Mining</i>	10
2.2.2	<i>Text Preprocessing</i>	10
2.2.3	<i>Case Folding</i>	10
2.2.4	<i>Tokenizing</i>	10
2.2.5	<i>Stop word Filtering</i>	11
2.2.6	<i>Stemming</i>	11
2.2.7	Algoritma	16
2.2.8	Algoritma Pencocokan <i>String</i>	17
2.2.9	ASCII	17
2.2.10	Algoritma Rabin Karp.....	19
BAB III ANALISIS DAN PERANCANGAN		22
3.1	Analisis Masalah	22
3.2	Hasil Analisis	22
3.3	Solusi Yang Dapat Di Terapkan.....	23
3.4	Analisa Kebutuhan	23
3.4.1	Kebutuhan Perangkat Keras (<i>Hardware</i>).....	24
3.4.2	Kebutuhan Perangkat Lunak (<i>Software</i>)	24
3.4.3	Kebutuhan Fungsional	25
3.4.4	Kebutuhan Non Fungsional.....	25
3.5	Analisa kebutuhan	26
3.5.1	Kelayakan Teknologi	26
3.5.2	Kelayakan Hukum.....	26
3.6	Perancangan Sistem.....	26
3.6.1	Perancangan Proses	27
BAB IV IMPLEMENTASI DAN PEMBAHASAN		35
4.1	<i>Database</i> dan Tabel	35
4.2	Antarmuka (<i>Interface</i>)	37

4.3	Koneksi <i>Database</i>	39
4.4	<i>White-Box Testing</i>	40
4.5	<i>Black-box Testing</i>	48
4.5.1	Pengujian Waktu Proses.....	49
4.5.2	Pengujian Akurasi	51
4.6	Implementasi Program	54
BAB V PENUTUP.....		55
5.1	Kesimpulan.....	55
5.2	Saran.....	55
DAFTAR PUSTAKA		57



DAFTAR TABEL

Tabel 2.1 Imbuhan Yang Tidak Diizinkan.....	13
Tabel 2.2 Aturan Penghapusan Awalan Me-	13
Tabel 2.3 Aturan Penghapusan Awalan Pe-.....	14
Tabel 2.4 Aturan Penghapusan Awalan Be-.....	15
Tabel 2.5 Aturan Penghapusan Awalan Te-	15
Tabel 4.1 Uji Waktu Proses Filtering Stop Word	50
Tabel 4.2 Pengujian Akurasi Algoritma Stemming Nazief & Adriani	53



DAFTAR GAMBAR

Gambar 2.1 Tabel ASCCI.....	19
Gambar 3.1 Use Case <i>User</i>	27
Gambar 3.2 Flowchat Keseluruahn Sistem.....	28
Gambar 3.3 Flowchart Preprocessing Text.....	28
Gambar 3.4 Flowchart Filtering.....	29
Gambar 3.5 Flowchart Filtering Stop Word.....	30
Gambar 3.6 Flowchart Algoritma Nazief & Adriani.....	31
Gambar 3.7 Struktur Tabel <i>tb_katadasar</i>	33
Gambar 3.8 Struktur Tabel <i>tb_stopword</i>	33
Gambar 3.9 Antarmuka (interface) aplikasi.....	34
Gambar 4.1 Import Database <i>tb_stopword</i>	36
Gambar 4.2 <i>Import Database</i> <i>tb_katadasar</i>	36
Gambar 4.3 Layout Aplikasi Berdasarkan Framework Css Bootstrap.....	37
Gambar 4.4 Layout Aplikasi.....	38
Gambar 4.5 Apache2 Dan Php Sudah Di <i>Install</i>	40
Gambar 4.6 Status Apache <i>Server</i> Dan Mysql	42
Gambar 4.7 <i>White-Box Testing</i> Fungsi <i>stopword_hilang</i>	43
Gambar 4.8 White-Box Testing Fungsi <i>kataDasar</i>	47
Gambar 4.9 Uji File “ <i>setmming-nofilter.php</i> ”	50
Gambar 4.10 Uji File “ <i>nostemming.php</i> ”	52
Gambar 4.11 Uji File “ <i>stemming.php</i> ”	52

INTISARI

Tingkat akurasi pada algoritma pencocokan kata bergantung pada pre-processing sehingga perlu diterapkan stemming. stemming Algoritma nazief & adriani pada pre-processing dapat menambah akurasi pada algoritma pencocokan kata. Penambahan stemming pada pre-processing dapat menambah waktu proses sebuah aplikasi, dibutuhkan filtering khusus untuk mengurangi kata-kata yang akan di stemming.

Berdasarkan permasalahan di atas peneliti menggunakan metode penelitian *Research and Development*. Metode ini digunakan untuk menghasilkan produk tertentu dan menguji keefektifan produk tersebut. Algoritma nazief & adriani memiliki waktu proses yang cukup lama, sehingga perlu penambahan filtering kata umum untuk mengurangi waktu proses pada algoritma.

Aplikasi yang dihasilkan berupa file “.php” yang ditunjukkan untuk memberikan gambaran atau hasil penelitian penerapan stemming dan filtering pada pre-processing sehingga dapat menambah akurasi dan mengurangi waktu proses pada saat stemming.

Kata-kunci: Algoritma, rabin karp, stemming nazief & adriani, filtering kata umum

ABSTRACT

Accuracy in word matching algorithms depends on pre-processing so it needs to be applied stemming. stemming Pre-processing nazief & adriani algorithms can add accuracy to word matching algorithms. Adding stemming to pre-processing can add to the processing time of an application, special filtering is needed to reduce the words that will be stemming.

Based on the above problems the researcher used Research and Development research methods. This method is used to produce certain products and test the effectiveness of these products. The nazief & adriani algorithm has a long processing time, so it is necessary to add filtering stop word to reduce the processing time in the algorithm.

The resulting application is in the form of a ".php" file which is shown to provide an overview or the results of research on the application of stemming and filtering in pre-processing so that it can increase accuracy and reduce processing time when stemming.

Keywords: *Algorithm, rabin karp, nazief & adriani stem, filtering stop word*