

*Stemming Bahasa Jerman Menggunakan Algoritma Comprehensive Morphological Rules Berbasis Corpus*

*Diajukan Untuk Menyusun Tugas Akhir di Jurusan Teknik Informatika  
Fakultas Ilmu Komputer Universitas Sriwijaya*



Oleh :

**DWI ERVIANA**

**09121002007**

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**

**2018**

# LEMBAR PENGESAHAN TUGAS AKHIR

## **STEMMING BAHASA JERMAN MENGGUNAKAN ALGORITMA *COMPREHENSIVE MORPHOLOGICAL RULES* BERBASIS *CORPUS***

Oleh :

**Dwi Erviana**

**NIM: 09121002007**

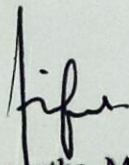
Palembang, Maret 2018

Pembimbing II,

**Novi Yusliani, M.T.**

**NIP 198211082012122001**

Pembimbing I



**Rifkie Primartha, M.T.**

**NIP. 197706012009121004**

Mengetahui,

Ketua Jurusan Teknik Informatika



**Rifkie Primartha, M.T.**  
**NIP: 197706012009121004**



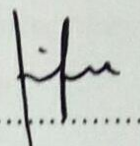
## TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Jumat tanggal 9 Maret 2018 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Dwi Erviana  
NIM : 09121002007  
Judul : *Stemming Bahasa Jerman Menggunakan Algoritma Comprehensive Morphological Rules Berbasis Corpus*

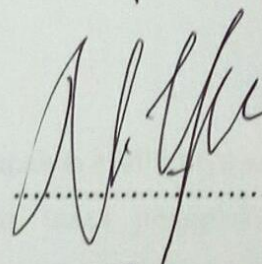
1. Ketua Penguji

Rifkie Primartha, M.T  
NIP. 197706012009121004



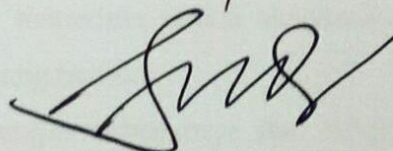
2. Sekretaris

Novi Yusliani, M.T  
NIP 198211082012122001



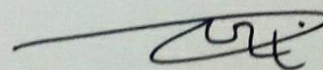
3. Penguji I

Rusdi Efendi, M.Kom  
NIP. 198201022011021201

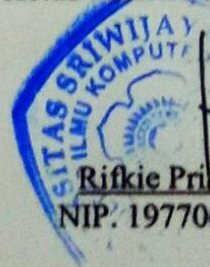


4. Penguji II

Osvari Arsalan, M.T  
NIP. 1601142806880003



Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T  
NIP. 197706012009121004



## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Dwi Erviana  
NIM : 09121002007  
Program Studi : Teknik Informatika  
Judul Skripsi : *Stemming Bahasa Jerman Menggunakan Algoritma Comprehensive Morphological Rules Berbasis Corpus*  
Hasil Pengecekan  
*Software iThenticate/Turnitin* : 8%

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, Maret 2018



DWI ERVIANA  
NIM. 09121002007

## Motto dan Persembahan

### Motto :

- “Barang siapa keluar untuk mencari ilmu maka dia berada di jalan Allah.”  
(HR Tarmidzi)
- “Pendidikan merupakan perlengkapan paling baik untuk hari tua.”  
(Aristoteles)
- Wisuda setelah 12 semester adalah kesuksesan yang tertunda.
- Lebih baik terlambat daripada tidak wisuda sama sekali.

### Karya tulis ini Kupersembahkan kepada :

- Allah SWT
- Nabi Muhammad SAW
- Kedua Orang Tua
- Keluarga Besar
- Teman-teman Teknik Informatika Reguler 2012, serta
- Almamater

# ***STEMMING BAHASA JERMAN MENGGUNAKAN ALGORITMA COMPREHENSIVE MORPHOLOGICAL RULES BERBASIS CORPUS***

**Oleh:  
Dwi Erviana  
09121002007**

## **ABSTRAK**

Bahasa Jerman merupakan salah satu bahasa Internasional yang populer dan masih banyak dipelajari. Salah satu hal yang mempermudah pembelajaran Bahasa Jerman yaitu dengan cara membuat suatu kalimat yang berimbuhan menjadi kata dasar. Oleh karena itu penelitian ini mengimplementasikan metode *stemming* CMR yang dapat memproses perubahan kalimat tersebut. Penelitian ini menggunakan data sebanyak 30105 kata dasar. Data ini akan dicocokkan dengan data uji sebanyak 2384 kata. Hasil penelitian ini membuktikan bahwa tingkat akurasi pada algoritma *stemming* CMR sebesar 99%. Hal ini dikarenakan semakin banyak data didalam *corpus* maka semakin besar tingkat akurasi yang didapat.

Kata Kunci : *Stemming, German Stemming, CMR, Corpus.*

# **GERMAN STEMMING USING CORPUS BASED COMPREHENSIVE MORPHOLOGICAL RULES ALGORITHM**

**Written By:  
Dwi Erviana  
09121002007**

## **ABSTRACT**

German language is one of the most popular international language and still widely studied. One of the things that make the learning process of German language easier is by changing the the inflected or derived sentence into root form. Therefore this research implements the CMR stemming method that can process the conversion of the sentence. This research uses data of 30105 root words. These data will be matched with the test data of 2384 words. The results of this research proves that the accuracy level of CMR stemming algorithm is 99%. This is because the more data in the corpus, the greater level of accuracy it can gained.

Keywords : *Stemming, German Stemming, CMR, Corpus.*

## KATA PENGANTAR

Puji dan syukur kehadiran Allah SWT karena atas rahmat-Nya penulis dapat menyelesaikan tugas akhir ini. Tugas akhir yang berjudul “***Stemming Bahasa Jerman Menggunakan Algoritma Comprehensive Morphological Rules Berbasis Corpus***” ini disusun untuk memenuhi salah satu persyaratan kelulusan tingkat S1 pada Jurusan Teknik Informatika Universitas Sriwijaya.

Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada pihak-pihak telah memberikan dukungan, bimbingan dan motivasi kepada penulis untuk menyelesaikan tugas akhir ini, yaitu kepada:

1. Kedua orang tua tercinta, terima kasih atas dukungan, semangat, kasih sayang, serta perjuangan dan kerja kerasnya yang tak kenal lelah guna merawat, menghidupi, dan membiayaiku. Terima kasih untuk selalu menyebut namaku dalam setiap doa.
2. Saudara-saudara dan keluargaku tersayang, Yarica Eryana, Muhammad Evrian, Sahri Putra Fernandes dan Nasya Batrisya Fernandes, terima kasih untuk semangat dan dukungannya.
3. Bapak Jaidan Jauhari, S.Pd., M.T. selaku dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Rifkie Primartha, M.T. selaku Ketua Jurusan Teknik Informatika Universitas Sriwijaya dan Pembimbing I untuk tugas akhir saya, terima kasih atas bimbingan dan ilmu yang diberikan selama menyelesaikan tugas akhir ini.
5. Ibu Novi Yusliani, M.T. selaku Pembimbing II yang telah banyak memberikan bimbingan, masukan dan bantuan terbaik kepada Penulis dalam menyelesaikan tugas akhir ini.
6. Bapak Rusdi Efendi, M.Kom. dan Bapak Osvari Arsalan, M.T. selaku Penguji Tugas Akhir, terima kasih atas nasihat dan saran yang diberikan.
7. Seluruh dosen yang telah memberikan ilmunya selama Penulis menuntut ilmu di Fasilkom Univeritas Sriwijaya.



8. Sahabat-sahabat Parjim Squad Dilla, Putri, Dara dan Wenty yang telah menjadi teman seperjuangan, menjadi sahabat, menjadi saudara dan selalu menemani dan membantu sejak awal perkuliahan.
9. Auzan Lazuardi yang telah mendukung dan banyak membantu penulis untuk dapat menyelesaikan skripsi ini.
10. Teman-teman PP Palembang-Indralaya Sarniko, Renaldy, Nanda, Hisbulah yang telah banyak membantu sejak awal perkuliahan, menjadi sahabat dan menjadi teman dalam susah dan senang mengejar TM dan bus kuning.
11. Seluruh mahasiswa Jurusan Teknik Informatika Angkatan 2012 dan seluruh teman-teman yang turut membantu dan mendukung Penulis dalam hal perkuliahan.
12. Admin jurusan Teknik Informatika, Mbak Winda dan seluruh staf administrasi serta pegawai yang selalu membantu dan mendukung penulis dalam hal administrasi perkuliahan.
13. Beserta semua pihak-pihak lain yang tidak dapat Penulis sebutkan satu persatu yang terlibat dalam penyelesaian Skripsi ini.

Terselesainya Tugas Akhir ini tentunya tidak lepas dari bantuan, bimbingan, dorongan, serta doa dari berbagai pihak. Untuk itu, Penulis mengucapkan rasa terima kasih yang sebesar-besarnya atas berbagai bantuannya tersebut. Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna. Oleh karena itu, berbagai bentuk kritik, saran dan koreksi sangat Penulis harapkan demi kesempurnaan Tugas Akhir ini dan semoga tugas akhir ini dapat bermanfaat bagi ilmu pengetahuan.

Akhir kata, penulis menyadari bahwa tugas akhir ini jauh dari kata sempurna. Untuk itu penulis mengharapkan kritik dan saran yang membangun dari semua pihak untuk penyempurnaan tugas akhir ini dan semoga tugas akhir ini dapat bermanfaat bagi pihak yang membutuhkan.

Palembang, Maret 2018

Penulis

## DAFTAR ISI

	Halaman
HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN .....	ii
HALAMAN TANDA LULUS UJIAN SIDANG TUGAS AKHIR.....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL .....	xiv
DAFTAR LAMPIRAN .....	xv
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	I-1
1.2 Rumusan Masalah .....	I-3
1.3 Tujuan Penelitian .....	I-3
1.4 Manfaat Penelitian .....	I-3
1.5 Batasan Masalah .....	I-3
1.6 Metodologi Penelitian .....	I-4
1.6.1 Jenis dan Sumber Data.....	I-4
1.6.2 Teknik Pengumpulan Data.....	I-4
1.6.3 Tahapan Penelitian.....	I-4
1.7 Metoda Pengembangan Perangkat Lunak.....	I-5
1.8 Sistematika Penulisan .....	I-8
<b>BAB II LANDASAN TEORI</b>	
2.1 Penelitian Terkait .....	II-1

2.2	Bahasa Jerman.....	II-2
2.3	<i>Stemming</i> .....	II-4
2.4	<i>Comprehensive Morphological Rules</i> .....	II-4
2.4.1	CMR Berbasis <i>Corpus</i> .....	II-6
2.4.2	CMR Berbasis <i>Non-Corpus</i> .....	II-8
2.5	Pra-pengolahan .....	II-9
2.6	<i>Rational Unified Process (RUP)</i> .....	II-11

### BAB III ANALISIS DAN PERENCANAAN

3.1	Analisis Masalah .....	III-1
3.1.1	Analisis Data .....	III-1
3.1.2	Analisis Arsitektur Perangkat Lunak .....	III-3
3.1.3	Analisis Proses Pra-pengolahan .....	III-4
3.1.3.1	<i>Case Folding</i> .....	III-5
3.1.3.2	<i>Tokenizing</i> .....	III-5
3.1.3.3	<i>Stopword Removal</i> .....	III-6
3.1.4	Analisis Algoritma <i>Stemming</i> <i>CMR Corpus</i> .....	III-7
3.2	Analisis Kebutuhan Perangkat Lunak .....	III-8
3.2.1	Deskripsi Umum Sistem .....	III-8
3.2.2	Spesifikasi Kebutuhan Perangkat Lunak .....	III-9
3.2.3	Model <i>Use Case</i> .....	III-9
3.2.3.1	Diagram <i>Use Case</i> .....	III-9
3.2.3.2	Definisi Aktor .....	III-10
3.2.3.3	Definisi <i>Use Case</i> .....	III-10
3.2.3.4	Skenario <i>Use Case</i> .....	III-11
3.3	Realisasi <i>Use Case</i> .....	III-12
3.3.1	Use Case Pra-pengolahan .....	III-12
3.3.1.1	Diagram Kelas Analisis Pra-pengolahan .....	III-12
3.3.1.2	Diagram <i>Sequence</i> Pra-pengolahan .....	III-13
3.3.2	<i>Use Case Stemming</i> Bahasa Jerman .....	III-14



3.3.2.1 Diagram Kelas Analisis <i>Stemming</i> Bahasa Jerman .....	III-14
3.3.2.2 Diagram <i>Sequence Stemming</i> Bahasa Jerman ..	III-15
3.4 Diagram <i>Class Stemming</i> Bahasa Jerman .....	III-16
3.5 Perancangan Antar Muka .....	III-17
<b>BAB IV IMPLEMENTASI DAN PENGUJIAN</b>	
4.1 Implementasi Perangkat Lunak .....	IV-1
4.1.1 Lingkungan Implementasi .....	IV-1
4.1.2 Implementasi Kelas .....	IV-2
4.1.3 Implementasi Antar Muka .....	IV-2
4.2 Pengujian Perangkat Lunak .....	IV-3
4.2.1 Lingkungan Pengujian .....	IV-3
4.2.2 Rencana Pengujian .....	IV-3
4.2.3 Kasus Uji .....	IV-5
4.2.4 Hasil Pengujian Perangkat Lunak .....	IV-7
4.3 Hasil, Analisis dan Eksperimen .....	IV-7
4.3.1 Analisis Akurasi .....	IV-8
4.3.1.1 Analisis Akurasi Algoritma <i>Stemming</i> CMR Berbasis <i>Corpus</i> .....	IV-9
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan .....	V-1
5.2 Saran .....	V-1
DAFTAR PUSTAKA .....	xvi

## DAFTAR GAMBAR

	Halaman
II-1. Arsitektur algoritma <i>stemming</i> bahasa Sunda berbasis <i>corpus</i> .....	II-7
II-2. Arsitektur algoritma <i>stemming</i> bahasa Jerman berbasis <i>non-corpus</i> ..	II-8
II-3. Ilustrasi proses <i>Case Folding</i> .....	II-10
II-4. Ilustrasi proses <i>Tokenizing</i> .....	II-10
II-5. Ilustrasi proses <i>Stopword Removal</i> .....	II-11
II-6. Arsitektur <i>Rational Unified Process</i> .....	II-13
III-1. Diagram proses algoritma <i>stemming</i> bahasa Jerman dengan pendekatan CMR berbasis <i>corpus</i> . .....	III-4
III-2. Diagram proses Pra-pengolahan.....	III-4
III-3. Contoh Proses <i>Case Folding</i> .....	III-5
III-4. Contoh Proses <i>Tokenizing</i> .....	III-5
III-5. Contoh Proses <i>Stopword Removal</i> .....	III-6
III-6. Contoh Proses Hapus Karakter Khusus .....	III-6
III-7. Arsitektur Algoritma <i>Stemming CMR Corpus</i> .....	III-7
III-8. Diagram <i>Use Case</i> .....	III-10
III-9. Diagram Kelas Analisis Pra-pengolahan.....	III-12
III-10. Diagram <i>Sequence</i> Pra-pengolahan .....	III-13
III-11. Diagram Kelas Analisis <i>Stemming</i> Bahasa Jerman .....	III-14
III-12. Diagram <i>Sequence Stemming</i> Bahasa Jerman .....	III-15
III-13. Diagram <i>Class Stemming</i> Bahasa Jerman .....	III-16
III-14. Rancangan Antar Muka .....	III-17
IV-1. Antarmuka <i>User</i> .....	IV-3
IV-2. Hasil Pengujian .....	IV-7

## DAFTAR TABEL

	Halaman
I-1. Kegiatan Pengembangan Perangkat Lunak Berdasarkan RUP .....	I-6
II-1. Contoh kata yang berimbuhan dalam bahasa Jerman .....	II-3
II-2. Contoh kata yang mengalami proses morfologis dalam bahasa Indonesia.....	II-5
II-3. Contoh kata yang mengalami proses morfologis dalam bahasa Jerman.....	II-6
III-1. Akhiran dalam Bahasa Jerman .....	III-1
III-2. Kebutuhan Fungsional Perangkat Lunak .....	III-9
III-3. Kebutuhan Non-Fungsional Perangkat Lunak .....	III-9
III-4. Definisi Aktor .....	III-10
III-5. Definisi <i>Use Case</i> .....	III-10
III-6. Skenario <i>Use Case</i> Pra-pengolahan .....	III-11
III-7. Skenario <i>Use Case Stemming</i> Bahasa Jerman .....	III-11
IV-1. Daftar Kelas .....	IV-2
IV-2. Rencana Pengujian <i>Use Case</i> Pra-pengolahan .....	IV-4
IV-3. Rencana Pengujian <i>Use Case Stemming</i> Bahasa Jerman .....	IV-4
IV-4. Kasus Uji <i>Use Case</i> Pra-pengolahan .....	IV-5
IV-5. Kasus Uji <i>Use Case Stemming</i> bahasa Jerman .....	IV-6
IV-6. Contoh Hasil Data Uji Algoritma <i>Stemming</i> CMR berbasis <i>Corpus</i> .	IV-7
IV-7. Analisis Hasil <i>Stemming</i> yang Gagal Diproses .....	IV-9
IV-8. Persentase Hasil Data Uji CMR Berbasis <i>Corpus</i> .....	IV-11



## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Hasil Pengujian .....	A-1
Lampiran 2. <i>Coding</i> Program.....	B-1

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Bahasa Jerman adalah salah satu bahasa yang digunakan di Eropa dan pernah menjadi bahasa pengantar antarbangsa sampai awal abad ke-20. Meskipun penggunaan bahasa Jerman di dunia saat ini sudah menurun, bahasa ini masih termasuk salah satu bahasa internasional yang populer dan masih banyak yang mempelajari karena banyak literatur klasik dunia yang ditulis menggunakan bahasa ini. Morfologi dari suatu bahasa dengan bahasa lainnya mempunyai perbedaan dalam tingkat kompleksitas (Agusta, 2009). Sehingga untuk proses *stemming* bahasa Jerman berbeda dengan proses *stemming* pada bahasa lainnya, seperti bahasa Indonesia.

*Stemming* merupakan proses penghilangan suatu kata yang berimbuhan menjadi kata dasar. Proses *stemming* ini merupakan hal penting dalam *Information Retrieval* (IR), IR akan menghasilkan dokumen yang efektif dan efisien dengan adanya proses *stemming* (Asian, Williams & Tahaghogi, 2005). Efektif berarti pengguna mendapatkan hasil yang relevan dengan masukan, dan efisien berarti waktu pencarian yang sesingkat-singkatnya (Agusta, 2009).

Penelitian *stemming* teks bahasa Indonesia berjudul “*Automatic Learning of Stemming Rules for the Indonesian Language*” menggunakan pendekatan *Comprehensive Morphological Rules* (CMR) berbasis *corpus* menunjukkan tingkat

akurasi sebesar 90% (Indradjaja, 2003). Proses *stemming* dokumen teks berbahasa Indonesia menggunakan algoritma Porter membutuhkan waktu yang lebih singkat dibandingkan dengan *stemming* menggunakan algoritma Nazief & Adriani (Agusta, 2009). Penelitian selanjutnya tentang *stemming* bahasa Jerman dilakukan oleh (Nicolai, 2016) yang berjudul “*Leveraging Inflection Tables for Stemming and Lemmatization*” dengan algoritma *stemming*. Hasil dari penelitian tersebut menunjukkan perbandingan tingkat akurasi *stemming* pada 3 bahasa, yaitu bahasa Inggris, bahasa Belanda dan bahasa Jerman.

*Comprehensive Morphological Rules (CMR)* merupakan aturan-aturan morfologi bahasa secara menyeluruh. Secara umum untuk menghasilkan kata dasar dalam proses *stemming* dengan algoritma CMR mempunyai dua cara, yaitu dengan algoritma CMR berbasis *corpus* dan algoritma CMR berbasis *non-corpus*. Proses dengan berbasis *corpus* yaitu melakukan pencarian ke dalam kumpulan suatu teks atau *database* (Asian, Williams & Tahaghogi, 2005). Contoh proses *stemming* berbasis *corpus* yaitu *stemming* bahasa Tamil yang diteliti oleh (Rajalingam, 2012). Proses *stemming* berbasis *corpus* memerlukan pengecekan ke dalam *database*, maka algoritma ini mempunyai tingkat akurasi yang tinggi, namun waktu proses yang dibutuhkan cukup lama. Sedangkan proses dengan berbasis *non-corpus* yaitu tanpa melakukan pencarian ke dalam kumpulan suatu teks atau *database*, sehingga waktu proses yang dibutuhkan lebih cepat, namun mempunyai tingkat akurasi yang lebih kecil dari algoritma berbasis *corpus* (Agusta, 2009). Contoh algoritma *stemming* yang berbasis *non-corpus* yaitu algoritma *stemming* Porter untuk teks bahasa Indonesia (Tala, 2003).



Berdasarkan uraian sebelumnya, terdapat kelebihan dari algoritma pendekatan CMR berbasis *non-corpus* pada *stemming* bahasa Indonesia dalam tingkat kecepatan, tetapi pada tingkat akurasi yang kecil. Sedangkan *stemming* berbasis *corpus* mempunyai tingkat akurasi yang tinggi. Karena itu, penelitian ini akan mengimplementasikan algoritma *stemming* dengan pendekatan CMR berbasis *corpus* pada bahasa Jerman untuk meningkatkan akurasi.

## **1.2 Rumusan Masalah**

Rumusan masalah pada penelitian ini adalah bagaimana tingkat akurasi dalam melakukan *stemming* teks bahasa Jerman menggunakan algoritma *Comprehensive Morphological Rules* berbasis *corpus*.

## **1.3 Tujuan Penelitian**

Tujuan penelitian ini adalah mengukur tingkat akurasi algoritma *stemming* pada perangkat lunak dengan menggunakan algoritma *Comprehensive Morphological Rules* berbasis *corpus* dalam *stemming* bahasa Jerman.

## **1.4 Manfaat Penelitian**

Manfaat penelitian ini dapat digunakan sebagai referensi penelitian selanjutnya dalam bidang *Information Retrieval (IR)* untuk bahasa Jerman.

## **1.5 Batasan Masalah**

Batasan masalah dari penelitian ini adalah :

1. Sistem hanya menerima masukan kata berbahasa Jerman sesuai dengan tata bahasa Jerman.
2. Sistem tidak dapat menerima masukan berupa singkatan kata.

## 1.6 Metodologi Penelitian

### 1.6.1 Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder berupa kumpulan kata dasar dalam bahasa Jerman. Terdapat satu sumber data yang dibutuhkan dalam penelitian ini, yaitu data dari website Institut für Deutsche Sprache (Institut Bahasa Jerman).

### 1.6.2 Teknik Pengumpulan Data

Teknik pengumpulan data yang dilakukan dalam penelitian ini adalah dengan cara memasukkan kata dasar kedalam corpus dengan meng-*konvert* kata dasar dari file (.txt) menjadi *database* dalam bentuk data (.csv).

### 1.6.3 Tahapan Penelitian

Tahapan yang dilakukan dalam penelitian algoritma *stemming* bahasa Jerman dengan pendekatan CMR berbasis *corpus* adalah :

1. Memahami algoritma *stemming* dengan pendekatan CMR berbasis *corpus*.
2. Membuat aturan-aturan imbuhan dalam bahasa Jerman untuk diterapkan ke dalam proses *stemming* berdasarkan pendekatan CMR *corpus* .
3. Membuat kamus data untuk kata dasar bahasa Jerman.
4. Melakukan pengembangan perangkat lunak dengan menggunakan metode *Rational Unified Process* (RUP).
5. Menganalisa dan membahas hasil pengujian perangkat lunak dan sistem.
6. Menarik kesimpulan dan menyelesaikan laporan penelitian.

## 1.7 Metode Pengembangan Perangkat Lunak

Metode yang diterapkan dalam pengembangan perangkat lunak ini adalah *Rational Unified Process* (RUP) yang merupakan model pengembangan perangkat lunak berorientasi objek dan bersifat *iterative incremental*. RUP dapat diadaptasi dan diperluas sehingga penggunaan metode ini dapat disesuaikan dengan kebutuhan. Penggunaan RUP pada penelitian ini didasarkan pada beberapa kelebihan metode ini, diantaranya adalah mendukung pengembangan berorientasi objek, memungkinkan untuk mengontrol perubahan-perubahan yang terjadi pada perangkat lunak selama proses pengembangannya secara sistematis, dan metode ini juga mendukung proses pengulangan dan penambahan-penambahan proses dalam pengembangan perangkat lunak (Kruchten, 2000). Kegiatan yang dilakukan pada setiap fase RUP dapat dilihat pada Tabel I-1.

Tabel I-1. Kegiatan Pengembangan Perangkat Lunak Berdasarkan RUP

	<b>Insepsi</b>	<b>Elaborasi</b>	<b>Konstruksi</b>	<b>Transisi</b>
<b>Pemodelan bisnis</b>	Pembuatan skenario dan <i>use case</i>	Perbaikan skenario dan <i>use case</i> jika diperlukan	Penyempurnaan skenario dan <i>use case</i>	
<b>Kebutuhan</b>	Pengumpulan kebutuhan dan batasan masalah	Analisis arsitektur perangkat lunak	Memastikan kembali kebutuhan dan batasan masalah	
<b>Analisis dan perancangan</b>	Pembuatan model kelas analisis dan perancangan prototipe antarmuka	Pembuatan model kelas analisis, diagram kelas, diagram sekuen, dan diagram aktivitas berdasarkan <i>use case</i> pada fase insepsi	Perbaikan model kelas analisis, diagram kelas, diagram sekuen, dan diagram aktivitas jika diperlukan	Membuat dokumentasi cetak biru perangkat lunak.
<b>Implementasi</b>		Pengkodean antarmuka dan kelas-kelas	Penyelesaian kode program	Perbaikan kode program jika diperlukan
<b>Pengujian</b>	Perencanaan pengujian	Pembuatan prosedur pengujian	Pelaksanaan pengujian	Evaluasi terhadap penerapan perangkat lunak
<b>Penerapan</b>		Merencanakan penerapan perangkat lunak	Penerapan perangkat lunak	

<b>Manajemen Proyek</b>	Inialisasi masalah dan ruang lingkup dari proyek perangkat lunak, mengatur penjadwalan aktifitas penelitian	Membuat daftar kegiatan pengembangan perangkat lunak	Mengeksekusi proyek perangkat lunak, manajemen iterasi yang dilakukan pada setiap fase	Menganalisa hasil penelitian dan menarik kesimpulan, persiapan fase <i>closing</i> proyek perangkat lunak
-------------------------	---	--	--	---

## **1.8 Sistematika Penulisan**

Sistematika penulisan pada tugas akhir ini adalah sebagai berikut:

### **1. Bab I Pendahuluan**

Pada bab ini dibahas mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, metodologi penelitian, metode pengembangan perangkat lunak dan sistematika penulisan.

### **2. Bab II Landasan Teori**

Pada bab ini berisi landasan dasar teori yang akan digunakan dalam melakukan analisis, perancangan dan implementasi tugas akhir yang dilakukan pada bab-bab selanjutnya.

### **3. Bab III Analisis dan Perancangan**

Pada bab menjelaskan tentang analisis dan desain antarmuka algoritma stemming Bahasa Jerman.

### **4. Bab IV Implementasi dan Pengujian**

Pada bab ini dibahas mengenai implementasi program, hasil eksekusi dan hasil pengujian.

### **5. Bab V Kesimpulan dan Saran**

Pada bab ini berisi kesimpulan hasil pengujian perangkat lunak yang didapat dan saran untuk penelitian selanjutnya.



## DAFTAR PUSTAKA

- Adriani, M., Asian, J., Nazief, B., & Williams, H. E. (2007). Stemming Indonesian : A Confix-Stripping Approach. <https://doi.org/10.1145/1316457.1316459>.
- Agusta, L., Kristen, U., & Wacana, S. (2009). Perbandingan Algoritma Stemming Porter Dengan Algoritma Nazief & Adriani Untuk Stemming Dokumen Teks Bahasa Indonesia. *Konferensi Nasional Sistem Dan Informatika 2009*, (KNS&I09-036), 196–201.
- Al-hashemi, R. (2010). Text Summarization Extraction System ( TSES ) Using Extracted Keywords, *I*(4), 164–168.
- Aljohani, A., & Mohd, M. (2014). Arabic-English Cross-Language Plagiarism Detection Using Winnowing Algorithm.
- Asian, J., Williams, H. E., & Tahaghoghi, S. M. M. (2005). Stemming Indonesian. *Conferences in Research and Practice in Information Technology Series*, 38, 307–314. <https://doi.org/10.1145/1316457.1316459>
- Frankfurter Allgemeine Zeitung*. (<http://www.faz.net>, diakses tanggal 17 Januari 2018)
- Handelsblatt*. (<http://www.handelsblatt.com>, diakses tanggal 27 Desember 2017)
- Indradjaja, S. L., & Bressan, S. (2003). Automatic Learning of Stemming Rules for the Indonesian Language. *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation*, 62–68. Retrieved from <http://aclweb.org/anthology/Y03-1007>

- Junaedi, D., Herlistiono, O., & Akbar, D. (2010). Stemmer For “Basa Sunda,” 275–278.
- Kruchten, P. (2001). *What Is the Rational Unified Process?*
- M.Suneetha, & Fatima, S. S. (2011). Corpus based Automatic Text Summarization System with HMM Tagger. *International Journal of Soft Computing & Engineering*, 118–123.
- Nicolai, G., & Kondrak, G. (2016). Leveraging Inflection Tables for Stemming and Lemmatization.
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program: Electronic Library and Information Systems*, 14(3), 130–137.  
<https://doi.org/10.1108/eb046814>
- Rajalingam, D. (2012). A Rule Based Iterative Affix Stripping Stemming Algorithm For Tamil, 28–34.
- Savoy, J. (2006). Light Stemming Approaches for the French , Portuguese , German and Hungarian Languages, (March), 1031–1035.
- Suparman, T. (2008). Proses Morfologi dalam bahasa Indonesia.
- Tala, F. Z. (2003). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. *M.Sc. Thesis, Appendix D*, pp, 39–46.