**SKRIPSI**

**JUDUL DALAM BAHASA INDONESIA**

***JUDUL DALAM BAHASA INGGRIS***

****

NAMA LENGKAP

NIM

**PROGRAM STUDI S1 ILMU KOMPUTER**

**DEPARTEMEN ILMU KOMPUTER DAN ELEKTRONIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS GADJAH MADA**

**YOGYAKARTA**

**2018**

**SKRIPSI**

**JUDUL DALAM BAHASA INDONESIA**

***JUDUL DALAM BAHASA INGGRIS***

Diajukan untuk memenuhi salah satu syarat memperoleh derajat

Sarjana Ilmu Komputer

****

NAMA

NIM

**PROGRAM STUDI S1 ILMU KOMPUTER**

**DEPARTEMEN ILMU KOMPUTER DAN ELEKTRONIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS GADJAH MADA**

**YOGYAKARTA**

**2018**

#### HALAMAN PENGESAHAN

**SKRIPSI**

**JUDUL DALAM BAHASA INDONESIA**

Telah dipersiapkan dan disusun oleh

NAMA

NIM

Telah dipertahankan di depan Tim Penguji

pada tanggal 17 April 2018

Susunan Tim Penguji

Nama Pembimbing Nama Ketua Pengujij

Pembimbing Ketua Penguji

Nama Anggota Penguji

Anggota Penguji

#### PERNYATAAN

Dengan ini saya menyatakan bahwa Skripsi 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 diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 4 April 2018

Nama Lengkap

#### HALAMAN MOTTO DAN PERSEMBAHAN

*“Nihan ta kramanikang manah, bhrānta lunghā swābhawanya, akweh inangĕnangĕnya, dadi prārthana, dadi sangṣaya, pinakāwaknya, hana pwa wwang ikang wĕnang humṛt manah, sira tika manggĕh amanggih sukha, mangke ring paraloka waneh”*

*(Sārasamuccaya, 81)*

*“Tad viddhi pranipātena paripraśnena sevayā,*

*Upadekṣyanti te jñānam jñāninas tattva-darsinah”*

(*Bhagavad Gītā IV.34*)

*“From Bego to Hero”*

*(Unknown)*

*Karya ini kupersembahkan kepada*

*Bapak, Ibu, Kakak dan keluarga besar tercinta*

*Rekan-rekan seperjuangan Ilmu Komputer UGM 2014*

*Segenap keluarga Himakom OTI semua angkatan*

*Keluarga KKN-PPM UGM BTL-10 2017*

*Semua Guru, Staf Pengajar, dan Dosen yang Telah Membimbing Penulis*

#### PRAKATA

Puji syukur kehadirat Ida Sang Hyang Widhi Wasa, Tuhan Yang Maha Esa atas limpah rahmat, karunia, serta petunjuk-Nya sehingga tugas akhir berupa penyusunan skripsi dengan judul “*Implementasi Modifikasi Enhanced Confix Stripping Stemmer Untuk Bahasa Jawa Kuno (Kawi)*” telah terselesaikan dengan baik. Karya tulis ini disusun untuk memenuhi persyaratan memperoleh gelar sarjana pada Program Studi Ilmu Komputer, Departemen Ilmu Komputer dan Elektronika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada.

Proses pembuatan karya tulis ini tidak luput dari bimbingan dan bantuan dari berbagai pihak. Untuk itu penulis mengucapkan terima kasih kepada:

1. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
2. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
3. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
4. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
5. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
6. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
7. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
8. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
9. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
10. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

Penulis menyadari bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan kelemahan, maka dari itu kritik dan saran yang membangun sangat disarankan untuk mendapatkan kesempurnaan tugas akhir. Semoga skripsi ini dapat bermanfaat bagi kita semua dan lebih khusus lagi penulis, pembaca, dan pergembangan dunia ilmu komputer pada umumnya.

Yogyakarta, 4 April 2018

Penulis

Komang Ary Tebuana

#### DAFTAR ISI

[HALAMAN PENGESAHAN i](#_Toc512813881)

[PERNYATAAN ii](#_Toc512813882)

[HALAMAN MOTTO DAN PERSEMBAHAN iii](#_Toc512813883)

[PRAKATA iv](#_Toc512813884)

[DAFTAR ISI vi](#_Toc512813885)

[DAFTAR GAMBAR viii](#_Toc512813886)

[DAFTAR TABEL x](#_Toc512813887)

[INTISARI xi](#_Toc512813888)

[ABSTRACT xii](#_Toc512813889)

[BAB I PENDAHULUAN 1](#_Toc512813890)

[1.1. Latar Belakang 1](#_Toc512813891)

[1.2. Rumusan Masalah 3](#_Toc512813892)

[1.3. Batasan Masalah 3](#_Toc512813893)

[1.4. Tujuan Penelitian 4](#_Toc512813894)

[1.5. Manfaat Penelitian 4](#_Toc512813895)

[1.6. Metode Penelitian 4](#_Toc512813896)

[1.7. Sistematika Penulisan 5](#_Toc512813897)

[BAB II TINJAUAN PUSTAKA 7](#_Toc512813898)

[BAB III LANDASAN TEORI 12](#_Toc512813899)

[3.1. Bahasa Kawi 12](#_Toc512813900)

[3.2. Teknik *Stemming* 21](#_Toc512813901)

[3.2.1. Algoritma *Stemming* Nazief dan Adriani 23](#_Toc512813902)

[3.2.2. Algoritma *Confix-Stripping Stemmer* 25](#_Toc512813903)

[3.2.3. Algoritma *Enhanced Confix Stripping (ECS) Stemmer* 27](#_Toc512813904)

[3.2.4. Algoritma Porter 31](#_Toc512813905)

[3.3. Kesalahan *Stemming* 33](#_Toc512813906)

[3.4. *Regular Expression* 34](#_Toc512813907)

[BAB IV ANALISIS DAN PERANCANGAN 36](#_Toc512813908)

[4.1. Analisis Permasalahan 36](#_Toc512813909)

[4.2. Rancangan Umum Sistem 37](#_Toc512813910)

[4.3. Data 39](#_Toc512813911)

[4.3.1. Data Kamus 39](#_Toc512813912)

[4.3.2. Data Kata Uji 40](#_Toc512813913)

[4.3.3. Data Kata Dasar (Kunci Jawaban) 40](#_Toc512813914)

[4.4. Praproses Data 41](#_Toc512813915)

[4.4.1. *Case Folding* 41](#_Toc512813916)

[4.4.2. Tokenisasi 42](#_Toc512813917)

[4.5. Perancangan *Stemmer* 42](#_Toc512813918)

[4.5.1. *Enhanced Confix Stripping Stemmer* 43](#_Toc512813919)

[4.5.2. *Enhanced Confix Stripping Stemmer* Tanpa Kamus 50](#_Toc512813920)

[4.5.3. Porter *Stemmer* 53](#_Toc512813921)

[4.6. Perancangan Evaluasi 56](#_Toc512813922)

[BAB V IMPLEMENTASI 61](#_Toc512813923)

[5.1. Lingkungan Implementasi 61](#_Toc512813924)

[5.2. Implementasi Pengumpulan Data 61](#_Toc512813925)

[5.2.1. Pengumpulan Data Kamus 61](#_Toc512813926)

[5.2.2. Pengumpulan Data Uji 62](#_Toc512813927)

[5.2.3. Pengumpulan Data Kata Kunci 63](#_Toc512813928)

[5.3. Implementasi Praproses Data 64](#_Toc512813929)

[5.3.1. Implementasi *Case Folding* 64](#_Toc512813930)

[5.3.2. Implementasi Tokenisasi 64](#_Toc512813931)

[5.4. Implementasi Algoritma *Enhanced Confix Stripping* 65](#_Toc512813932)

[5.4.1. Pengecekan Kamus 66](#_Toc512813933)

[5.4.2. Penerapan Aturan Pemotongan Akhiran 68](#_Toc512813934)

[5.4.3. Penerapan Aturan Pemotongan Awalan dan Sisipan 69](#_Toc512813935)

[5.4.4. Penerapan *loopPengembalianAkhiran* 71](#_Toc512813936)

[5.4.5. Penerapan Alur Proses Algoritma 73](#_Toc512813937)

[5.5. Implementasi Algoritma Porter 74](#_Toc512813938)

[5.6. Implementasi Pengujian dan Evaluasi 77](#_Toc512813939)

[5.6.1. Evaluasi Waktu Proses 78](#_Toc512813940)

[5.6.2. Evaluasi Akurasi 78](#_Toc512813941)

[BAB VI HASIL DAN PEMBAHASAN 80](#_Toc512813942)

[6.1. Pengujian Sistem dan Hasil 80](#_Toc512813943)

[6.2. Hasil Praproses Data 80](#_Toc512813944)

[6.2.1. Hasil *Case Folding* 80](#_Toc512813945)

[6.2.2. Hasil Tokenisasi 81](#_Toc512813946)

[6.3. Hasil Proses *Stemming* 81](#_Toc512813947)

[6.4. Hasil Evaluasi Waktu Proses 82](#_Toc512813948)

[6.5. Hasil Evaluasi Akurasi 84](#_Toc512813949)

[6.5.1. Evaluasi Kesalahan *Stemming ECS* 86](#_Toc512813950)

[6.5.2. Evaluasi Kesalahan *Stemming ECS* tanpa Kamus dan Porter 89](#_Toc512813951)

[6.6. Komparasi Hasil Evaluasi Waktu Proses dengan Akurasi 91](#_Toc512813952)

[BAB VII KESIMPULAN 93](#_Toc512813953)

[7.1. Kesimpulan 93](#_Toc512813954)

[7.2. Saran 93](#_Toc512813955)

[DAFTAR PUSTAKA 95](#_Toc512813956)

[LAMPIRAN 97](#_Toc512813957)

#### DAFTAR GAMBAR

Gambar 3.1 Ilustrasi Proses *Stemming* (Purwoko, 2011) 21

Gambar 3.2 Metode *Conflation* (Sharma, 2012) 22

Gambar 3.3 Tipe dari Algoritma *Stemming* (Jivani, 2011) 23

Gambar 3.4 Alur Proses Porter (Tala, 2003) 31

Gambar 4.1 Diagram Alir Proses Sistem (1) 37

Gambar 4.2 Diagram Alir Proses Sistem (2) 38

Gambar 4.3 *Pseudocode Binary Search (iterative)* 39

Gambar 4.4 Cuplikan Tabel Verifikasi Kata Kunci 41

Gambar 4.5 Diagram Alir Praproses Data 41

Gambar 4.6 Diagram Alir Proses *Case Folding* 42

Gambar 4.7 Diagram Alir Proses Tokenisasi 42

Gambar 4.8 *Pseudocode* *Enhanced Confix Stripping Stemmer* (1) 45

Gambar 4.9 *Pseudocode* *Enhanced Confix Stripping Stemmer* (2) 46

Gambar 4.10 Diagram Alir *Stemming ECS* (1) 46

Gambar 4.11 Diagram Alir *Stemming ECS* (2) 47

Gambar 4.12 Diagram Alir *Stemming ECS* (3) 48

Gambar 4.13 *Pseudocode ECS Stemmer* Tanpa Kamus(1) 50

Gambar 4.14 *Pseudocode ECS Stemmer* Tanpa Kamus(2) 51

Gambar 4.15 Diagram Alir *Stemming ECS* Tanpa Kamus(1) 51

Gambar 4.16 Diagram Alir *Stemming ECS* Tanpa Kamus(2) 52

Gambar 4.17 *Pseudocode* Porter *Stemmer* 54

Gambar 4.18 Diagram AlirPorter *Stemmer* (1) 55

Gambar 4.19 Diagram AlirPorter *Stemmer* (2) 56

Gambar 4.20 *Confusion Matrix* (Fawcett, 2006) 58

Gambar 5.1 Kode Program Pengumpulan Data dengan *Scraping* 62

Gambar 5.2 Cuplikan Data Uji 63

Gambar 5.3 Cuplikan Hasil Validasi Kata Kunci 63

Gambar 5.4 Implementasi *Case Folding* 64

Gambar 5.5 Cuplikan Implementasi Tokenisasi (1) 64

Gambar 5.6 Cuplikan Implementasi Tokenisasi (2) 65

Gambar 5.7 Bentuk Data Kata Uji Setelah Tokenisasi 65

Gambar 5.8 Kode Program Pengecekan *Singular* atau *Plural* 66

Gambar 5.9 Kode Program Pengambilan File Kamus 67

Gambar 5.10 Cuplikan Method Pengecekan Kata pada Kamus 68

Gambar 5.11 Cuplikan Kode Program Pemotongan Akhiran (partikel) 68

Gambar 5.12 Cuplikan Kode Program Pemotongan Awalan (1) 69

Gambar 5.13 Cuplikan Kode Program Pemotongan Awalan (2) 70

Gambar 5.14 Kode Program *class AbstractDisambiguatePrefixRule* 71

Gambar 5.15 Kode Program Pemotongan Sisipan 71

Gambar 5.16 Kode Program *loopPengembalian Akhiran* (1) 72

Gambar 5.17 Kode Program *loopPengembalian Akhiran* (2) 73

Gambar 5.18 Cuplikan Kode Program Tahapan *Stemming ECS* (1) 73

Gambar 5.19 Cuplikan Kode Program Tahapan *Stemming ECS* (2) 74

Gambar 5.20 Kode Program *class DontStemShortWord* 74

Gambar 5.21 Cuplikan Kode Program *Stemming* Porter (1) 75

Gambar 5.22 Cuplikan Kode Program *Stemming* Porter (2) 76

Gambar 5.23 Cuplikan Kode Program *Stemming* Porter (3) 77

Gambar 5.24 Format Argumen Pengujian *Stemmer* 77

Gambar 5.25 Kode Program Implementasi Pengujian *ECS* *Stemmer* (1) 77

Gambar 5.26 Kode Program Implementasi Pengujian *ECS* *Stemmer* (2) 78

Gambar 5.27 Formula Ms. Excel Pengecekan Kecocokan Kata 79

Gambar 5.28 Cuplikan Tabel Evaluasi Akurasi 79

Gambar 5.29 Formula Ms. Excel Pengecekan Kesalahan *Stemming* 79

Gambar 6.1 Cuplikan Data Hasil Proses *Case Folding* 80

Gambar 6.2 Hasil Proses Tokenisasi 81

Gambar 6.3 Cuplikan Hasil Keluaran dari *Stemmer* 81

Gambar 6.4 Grafik Rata-rata Waktu Proses *Stemmer* 83

Gambar 6.5 Grafik Nilai Akurasi *Stemmer* 85

Gambar 6.6 Grafik Jumlah Kesalahan *Stemmer ECS* 87

Gambar 6.7 Diagram Evaluasi *Overstemming* dan *Understemming ECS* 89

Gambar 6.8 Diagram Evaluasi Tingkat Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter 90

Gambar 6.9 Diagram Komparasi Hasil Evaluasi *Stemmer* 91

#### DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka (1) 9

Tabel 2.3 Tinjauan Pustaka (2) 10

Tabel 2.3 Tinjauan Pustaka (3) 11

Tabel 3.1 Kombinasi Imbuhan Terlarang dalam Bahasa Indonesia 29

Tabel 3.2 Aturan Pemenggalan Awalan dalam Bahasa Indonesia 29

Tabel 3.3 Revisi Aturan Pemenggalan 30

Tabel 3.4 Aturan Untuk *Inflectional Particle* 32

Tabel 3.5 Aturan Untuk *Inflectional Possesive Pronoun* 32

Tabel 3.6 Aturan Untuk *First Order Derivational Prefix* 32

Tabel 3.7 Aturan Untuk *Second Order Derivational Prefix* 32

Tabel 3.8 Aturan Untuk *Derivational Suffix* 33

Tabel 3.9 Pola Umum Pada *Regex* 34

Tabel 3.10 Metode Pencocokan dan Pencarian *String* 35

Tabel 3.11 Metode Modifikasi *String* 35

Tabel 4.1 Tabel Pemenggalan Awalan *Stemmer* *ECS* 49

Tabel 4.2 Aturan Pemenggalan untuk Porter *Stemmer* 53

Tabel 4.3 Rancangan Analisis Kesalahan *Stemming* 60

Tabel 6.1 Hasil Evaluasi Waktu Proses *Stemmer* 82

Tabel 6.2 Hasil Evaluasi Akurasi *Stemmer* 84

Tabel 6.3 Evaluasi Kesalahan Hasil *Stemming ECS* 86

Tabel 6.4 Evaluasi Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter (1) 89

Tabel 6.5 Evaluasi Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter (2) 90

#### INTISARI

**IMPLEMENTASI MODIFIKASI ENHANCED CONFIX STRIPPING STEMMER UNTUK BAHASA JAWA KUNO (KAWI)**

Oleh

Komang Ary Tebuana

14/363864/PA/15866

Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

*Stemmer* Bahasa Kawi yang dibangun dalam penelitian ini menggunakan algoritma *Enhanced Confix Stripping*. Algoritma berbasis kamus tersebut diuji pula performanya apabila tanpa pengecekan kamus dan dengan beberapa metode pengecualian aturan pemotongan imbuhan. Selain itu, dibandingkan pula dengan *stemmer* yang dibangun dengan algoritma Porter. Obyek penelitian yang digunakan adalah kata-kata dalam Bahasa Kawi pada kitab Sarasamuscaya yang berjumlah 860 kata. Data uji diproses dengan menerapkan aturan pembentukan kata berimbuhan (*wangun*) dalam Bahasa Kawi tanpa aturan sandhi pada masing-masing algoritma *stemming*.

Keluaran dari *stemmer* yaitu berupa kata dasar hasil pemrosesan pemotongan imbuhan. Dari penelitian ini, didapatkan bahwa waktu proses dan akurasi bergantung pada metode yang digunakan. *Enhanced Confix Stripping* memperoleh nilai akurasi yang tertinggi, sedangkan *Enhanced Confix Stripping* tanpa kamus dengan waktu proses yang paling lambat dan terjadi penurunan akurasi. Dengan dilakukan pengecualian pemotongan pada imbuhan tertentu mampu mendapatkan peningkatan performa pada algoritma *ECS* tanpa kamus dan Porter untuk Bahasa Kawi.

**Kata Kunci:** *Stemming, Bahasa Kawi, Bahasa Jawa Kuno, Enhanced Confix Stripping, Porter.*

#### ABSTRACT

**IMPLEMENTATION OF MODIFICATION ENHANCED CONFIX STRIPPING STEMMER FOR OLD JAVANESE LANGUAGE (KAWI)**

By

Komang Ary Tebuana

14/363864/PA/15866

Digital processing in searching information on Kawi’s (Old Javanese) language document required filtering by using technique in information retrieval that is *Stemming*. Stemming used to find the root word that exists in a term or word with affixes by extracting words from affixes attached to the word.

The Kawi language *stemmer* built in this study uses *Enhanced Confix Stripping* algorithm. This dictionary-basef algorithm also tested for performance if without dictionary checking and with some exception method of affix cutting rule. In addition, this *stemmer* compared to stemmer build with Porter algorithm. The research object is words with Kawi language in Sarasamuscaya book which amounted to 860 words. The test data is processed by applying the formation (*wangun*) rules on each *stemming* algorithm.

The output of *stemmer* is the root word that the result of affix cutting process. From this research, it is found that process time and accuracy depend on the method used. *Enhanced Confix Stripping* receives the highest accuracy value, while *Enhanced Confix Stripping* without dictionary with the slowest processing time and decreased accuracy. With the exception of cutting at certain affixes it is able to get a performance improvement on the *stemming* algorithm without the dictionary for Kawi Language.

**Keyword:** *Stemming, Kawi Language, Old Javanese Language, Enhanced Confix Stripping, Porter.*

# PENDAHULUAN

## Latar Belakang

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ante massa, interdum in suscipit quis, laoreet non dolor. Aenean ut lectus justo. Sed metus magna, maximus et leo quis, ultricies facilisis sapien. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Ut sapien leo, iaculis in tincidunt ut, dignissim vitae orci. Vestibulum vitae blandit libero, vitae volutpat turpis. Ut luctus massa feugiat urna euismod porta. Fusce dapibus molestie dolor vulputate laoreet. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

## Rumusan Masalah

Dari uraian permasalahan di latar belakang, ….

## Batasan Masalah

Dalam penelitian ini, batasan masalah yang digunakan yaitu sebagai berikut.

1. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
2. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
3. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

## Tujuan Penelitian

Penelitian ini bertujuan untuk ....

## Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
2. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
3. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

## Metode Penelitian

Adapun metode yang digunakan dalam proses penelitian ini adalah sebagai berikut.

1. Studi Literatur

Pengumpulan bahan pustaka maupun literatur baik dari buku, jurnal, e-*book*, atau karya ilmiah yang relevan dengan topik Bahasa Kawi (Jawa Kuno) dan tata bahasanya, Teknik-teknik *Stemming*, dan algoritma *Enhanced Confix Stripping*.

1. Pengumpulan Data

Data yang dikumpulkan yaitu data kata uji yang diperoleh dari kitab Sarasamuscaya yang diekstrasi secara manual kemudian divalidasi oleh ahli untuk dibuat sebagai data kata kunci. Selain itu terdapat data kamus kata dasar yang didapatkan melalui sumber kamus Bahasa Kawi *online.*

1. Perancangan Sistem

Perancangan sistem dari *stemmer* untuk Bahasa Kawi dengan menggunakan algoritma *stemming* yaitu *Enhanced Confix Stripping Stemmer* yang disesuaikan dengan metode-metode yang akan diuji dan termasuk didalamnya perancangan normalisasi pada data. Perancangan meliputi pembuatan *flowchart* dari masing-masing metode pemrosesan dan pengujian.

1. Implementasi

Implementasi dilakukan dengan menggunakan *tools* yang telah disesuaikan pada analisis sistem dan perancangannya. Tahapan ini meliputi persiapan *environment*, praproses terhadap dataset, melakukan penerapan aturan *wangun* ke dalam algoritma *stemmer*, serta penyesuaian terhadap parameter dari *Enhanced Confix Stripping* agar memperoleh hasil yang maksimal.

1. Pengujian

Proses pengujian dilakukan dengan menghitung waktu proses dari *stemmer* dan tingkat akurasi (ketelitian) yang diperoleh dari hasil *stemming*.

## Sistematika Penulisan

Penelitian ini terdiri dari tujuh bab dengan sistematika masing-masing bab adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini memuat latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metodologi yang digunakan, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini memuat beberapa penelitian terdahulu yang terkait pada topik permasalahan, metode yang digunakan, dan menjadi bahan referensi dalam penelitian ini, serta untuk mengetahui posisi penelitian yang dilakukan terhadap penelitian sebelumnya.

BAB III LANDASAN TEORI

Bab ini memuat teori-teori yang digunakan dalam mendukung proses penelitian ini.

BAB IV ANALISIS DAN PERANCANGAN

Bab ini memuat penjelasan serta analisis permasalahan, sistem, data, dan perancangan terhadap sistem yang akan dibuat, dari arsitektur yang digunakan, algoritma, serta metode pengujiannya.

BAB V IMPLEMENTASI

Bab ini memuat spesifikasi *hardware* dan *software* yang digunakan dan hasil implementasi kode sistem yang dikembangkan berdasarkan perancangan yang dilakukan beserta penjelasannya.

BAB VI HASIL DAN PEMBAHASAN

Bab ini memuat rangkuman hasil penelitian dan pengujian berupa waktu proses, akurasi, dan permasalahan yang diperoleh beserta penjelasannya.

BAB V KESIMPULAN

Bab ini berisi kesimpulan penelitian yang telah dilakukan disertai saran untuk pengembangan penelitian selanjutnya.

# TINJAUAN PUSTAKA

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ante massa, interdum in suscipit quis, laoreet non dolor. Aenean ut lectus justo. Sed metus magna, maximus et leo quis, ultricies facilisis sapien. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Ut sapien leo, iaculis in tincidunt ut, dignissim vitae orci. Vestibulum vitae blandit libero, vitae volutpat turpis. Ut luctus massa feugiat urna euismod porta. Fusce dapibus molestie dolor vulputate laoreet. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

# LANDASAN TEORI

## Subbab

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ante massa, interdum in suscipit quis, laoreet non dolor. Aenean ut lectus justo. Sed metus magna, maximus et leo quis, ultricies facilisis sapien. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Ut sapien leo, iaculis in tincidunt ut, dignissim vitae orci. Vestibulum vitae blandit libero, vitae volutpat turpis. Ut luctus massa feugiat urna euismod porta. Fusce dapibus molestie dolor vulputate laoreet. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

**3.1.1. Anak Subbab**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ante massa, interdum in suscipit quis, laoreet non dolor. Aenean ut lectus justo. Sed metus magna, maximus et leo quis, ultricies facilisis sapien. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Ut sapien leo, iaculis in tincidunt ut, dignissim vitae orci. Vestibulum vitae blandit libero, vitae volutpat turpis. Ut luctus massa feugiat urna euismod porta. Fusce dapibus molestie dolor vulputate laoreet. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

**Anak-anak subbab**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ante massa, interdum in suscipit quis, laoreet non dolor. Aenean ut lectus justo. Sed metus magna, maximus et leo quis, ultricies facilisis sapien. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Ut sapien leo, iaculis in tincidunt ut, dignissim vitae orci. Vestibulum vitae blandit libero, vitae volutpat turpis. Ut luctus massa feugiat urna euismod porta. Fusce dapibus molestie dolor vulputate laoreet. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

Berdasarkan pengklasifikasian imbuhan-imbuhan di atas, maka bentuk kata berimbuhan dalam Bahasa Indonesia dapat dimodelkan sebagai berikut.

|  |  |
| --- | --- |
|  | (3.1) |

Algoritma yang dibuat oleh Bobby Nazief dan Mirna Adriani ini memiliki tahap-tahap sebagai berikut:

* 1. Cari kata yang akan di-*stem* dalam kamus. Jika ditemukan maka diasumsikan bahwa kata tesebut adalah *root word*. Maka algoritma berhenti.
  2. *Inflection Suffixes* (“-lah”, “-kah”, “-ku”, “-mu”, atau “-nya”) dibuang. Jika berupa *particles* (“-lah”, “-kah”, “-tah” atau “-pun”) maka langkah ini diulangi lagi untuk menghapus *Possesive Pronouns* (“-ku”, “-mu”, atau “-nya”), jika ada.
  3. Hapus *Derivation Suffixes* (“-i”, “-an” atau “-kan”). Jika kata ditemukan di kamus, maka algoritma berhenti. Jika tidak maka ke langkah 3a
     1. Jika “-an” telah dihapus dan huruf terakhir dari kata tersebut adalah “-k”, maka “-k” juga ikut dihapus. Jika kata tersebut ditemukan dalam kamus maka algoritma berhenti. Jika tidak ditemukan maka lakukan langkah 3b.
     2. Akhiran yang dihapus (“-i”, “-an” atau “-kan”) dikembalikan, lanjut ke langkah 4.
  4. Hapus *Derivation Prefix*. Jika pada langkah 3 ada sufiks yang dihapus maka pergi ke langkah 4a, jika tidak pergi ke langkah 4b.
     1. Periksa tabel kombinasi awalan-akhiran yang tidak diijinkan. Jika ditemukan maka algoritma berhenti, jika tidak pergi ke langkah 4b.
     2. For i = 1 to 3, tentukan tipe awalan kemudian hapus awalan. Jika *root word* belum juga ditemukan lakukan langkah 5, jika sudah maka algoritma berhenti. Catatan: jika awalan kedua sama dengan awalan pertama algoritma berhenti.
  5. Melakukan *Recoding*.
  6. Jika semua langkah telah selesai tetapi tidak juga berhasil maka kata awal diasumsikan sebagai *root word*. Proses selesai.

### Anak subbab

*Confix stripping* (CS) *stemmer* adalah sebuah metode *stemming* pada teks berbahasa Indonesia yang diperkenalkan oleh Jelita Asian (2005), dimana metode *Confix stripping* (CS) *stemmer* merupakan pengembangan dari metode *stemming* sebelumnya yang dibuat oleh Nazief dan Adriani (1996).

Langkah-langkah dari proses kerja algoritma *Confix Stripping Stemmer* adalah sebagai berikut (Adriani et al., 2007).

1. Kata yang hendak di-*stemming* dicari terlebih dahulu pada kamus. Jika ditemukan, berarti kata tersebut adalah kata dasar, jika tidak maka melakukan langkah yang kedua.
2. Cek *rule precedence*. Apabila suatu kata memiliki pasangan awalan-akhiran “be-lah”, “be-an”, “me-i”, “di-i”, “pe-i”, atau “te-i” maka langkah *stemming* selanjutnya adalah (5, 6, 3, 4, 7). Apabila kata tidak memiliki pasangan awalan-akhiran tersebut, langkah *stemming* berjalan normal (3, 4, 5, 6, 7).
3. Hilangkan *inflectional particle* P (“-lah”, “-kah”, “-tah”, “-pun”) dan kata ganti kepunyaan atau *possessive pronoun* PP (“-ku”, “-mu”, “-nya”).
4. Hilangkan *Derivation Suffixes* DS (“-i”, “-kan”, atau “-an”).
5. Hilangkan *Derivational Prefixes* DP {“di-”,“ke-”,“se-”,“me-”,“be-”,“pe”, “te-”} dengan iterasi maksimum adalah 3 kali:
   1. Langkah 5 ini berhenti jika:
      * Terjadi kombinasi imbuhan terlarang seperti pada Tabel 3.1.
      * Awalan yang dideteksi saat ini sama dengan awalan yang dihilangkan sebelumnya.
      * Tiga awalan telah dihilangkan.
   2. Identifikasikan tipe awalan dan hilangkan. Awalan ada dua tipe:
      * Standar: “di-”, “ke-”, “se-” yang dapat langsung dihilangkan dari kata.
      * Kompleks: “me-”, “be-”, “pe”, “te-” adalah tipe-tipe awalan yang dapat bermorfologi sesuai kata dasar yang mengikutinya. Oleh karena itu, gunakan aturan pada Tabel 3.2 untuk mendapatkan pemenggalan yang tepat.
   3. Cari kata yang telah dihilangkan awalannya ini di dalam kamus. Apabila tidak ditemukan, maka langkah 5 diulangi kembali. Apabila ditemukan, maka keseluruhan proses dihentikan.
6. Apabila setelah langkah 5 kata dasar masih belum ditemukan, maka proses recoding dilakukan dengan mengacu pada aturan pada Tabel 3.2. *Recoding* dilakukan dengan menambahkan karakter *recoding* di awal kata yang dipenggal. Pada Tabel 3.2, karakter *recoding* adalah karakter setelah tanda hubung (’-’) dan terkadang berada sebelum tanda kurung. Sebagai contoh, pada kata “menangkap” (aturan 15), setelah dipenggal menjadi “nangkap”. Karena tidak valid, maka *recoding* dilakukan dan menghasilkan kata “tangkap”.
7. Jika semua langkah gagal, maka input kata yang diuji pada algoritma ini dianggap sebagai kata dasar.

Apabila pada kata yang hendak di-*stemming* ditemukan tanda hubung (’-’), maka kemungkinan kata yang hendak di-*stemming* adalah kata ulang. *Stemming* untuk kata ulang dilakukan dengan memecah kata menjadi dua bagian yakni bagian kiri dan kanan (berdasarkan posisi tanda hubung ’-’) dan lakukan *stemming* (langkah 1-7) pada dua kata tersebut. Apabila hasil *stemming* keduanya sama, maka kata dasar berhasil didapatkan.

# ANALISIS DAN PERANCANGAN

## Analisis Permasalahan

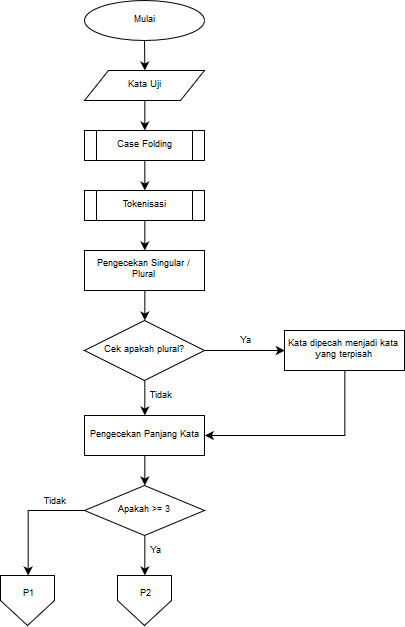
*Stemming* pada penelitian ini dilakukan untuk mengeksplorasi informasi serta guna kepentingan pengolahan digital pada dokumen-dokumen berbahasa Kawi. Belum terdapat *stemmer* yang digunakan untuk Bahasa Kawi sehingga untuk dibangun sebuah *stemmer* yang perlu diperhatikan adalah bentuk dari morfologi dari bahasanya. Dengan kemiripan bentuk morfologi antara Bahasa Indonesia dengan Bahasa Kawi maka akan diuji salah satu algoritma yaitu *Enhanced Confix Stripping,* dimana algoritma ini bekerja dengan menggunakan bantuan kamus sebagai referensi pengecekan katanya. Objek pengujian yang digunakan yaitu kata-kata Bahasa Kawi yang diperoleh dari kitab Sarasamuscaya, dimana kata yang tercantum sudah berbentuk huruf Latin dengan penyesuaian karakter huruf Kawi.

Algoritma *Enhanced Confix Stripping* yang dikembangkan untuk Bahasa Indonesia ini akan diuji dengan melakukan modifikasi pada aturan yang digunakan. Bagian yang menjadi pembeda adalah adanya tambahan fungsi penghapusan sisipan (infiks) yang tidak ada sebelumnya agar sesuai dengan aturan imbuhan dalam Bahasa Kawi yaitu awalan (*pengater*), sisipan (*seselan*), maupun akhiran (*pengiring*) yang terdapat pada Kamus Kawi-Bali dan tambahan aturan dari Miswanto (2012) sehingga dapat digunakan sebagai *stemmer* Bahasa Kawi. Dalam penelitian ini aturan pemenggalan kata yang dibangun untuk *stemmer* hanya akan menggunakan aturan *wangun* yang mencakup aturan *Nasal* pada beberapa imbuhandan tidak menggunakan aturan *sandhi*.

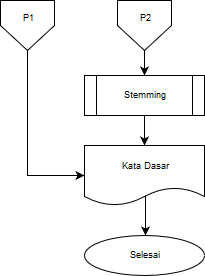
Pengujian performa dari algoritma *Enhanced Confix Stripping* pada Bahasa Kawi dengan beberapa metode pendekatan serta kondisi tertentu. Pengujian yang dilakukan yaitu dengan mengukur waktu proses kerja dari *stemmer* untuk mengetahui seberapa lama waktu yang dibutuhkan dalam melakukan *stemming* pada masing-masing pengujian serta tingkat akurasi dari kecocokan hasil *stemming* dengan kata dasarnya. Pengujian hasil *stemming* juga akan dikomparasikan dengan hasil dari implementasi algoritma Porter yang menggunakan aturan pemenggalan imbuhan Bahasa Kawi. Hasil dari pengujian selanjutnya akan dibandingkan untuk melihat performa dari masing-masing metode yang digunakan.

## Rancangan Umum Sistem

Pada penelitian ini dilakukan pembuatan *stemmer* untuk proses *stemming* teks yang menggunakan Bahasa Kawi. Adapun alur proses dari sistem yang dibangun dapat dilihat pada Gambar 4.1 s.d. Gambar 4.2.



**Gambar 4.1 Diagram Alir Proses Sistem (1)**



**Gambar 4.2 Diagram Alir Proses Sistem (2)**

Sistem yang dibuat terdiri dari beberapa langkah yaitu praproses (*case folding* dan tokenisasi) untuk menormalisasikan data masukan dan proses *stemming* dengan menggunakan algoritma *Enhanced Confix Stripping* dengan kamus dan tanpa kamus serta algoritma Porter. Sistem dibentuk berdasarkan aturan pemenggalan imbuhan dari masing-masing algoritma dimana *Enhanced Confix Stripping Stemmer* menggunakan *regular expression* untuk pemotongan imbuhannya sedangkan algoritma Porter lebih dominan menggunakan *affix removal*.

Proses kerja dari *stemmer* yang dibangun pada dasarnya memiliki alur kerja yang sama. Setiap data yang telah dinormalisasi akan langsung diproses oleh *stemmer*. Kata uji yang masuk akan dilakukan pengecekan terlebih dahulu apakah kata tergolong dalam jenis kata *singular* atau kata *plural*. Untuk kata yang *singular* (tunggal) maka akan masuk ke proses *stemming* dengan mengikuti *rule* yang digunakan untuk selanjutnya menghilangkan imbuhan-imbuhan yang bersesuaian dengan kata yang diuji dan menghasilkan keluaran berupa kata dasar. Sebelumnya akan dicek terlebih dahulu oleh sistem panjang dari kata tersebut, jika panjang kata ≤ 3 maka kata tidak akan diproses. Sedangkan untuk kata yang *plural* (berulang) akan dipecah menjadi kata yang terpisah yang selanjutnya diproses setiap katanya untuk dihilangkan imbuhannya sehingga akan menghasilkan kata dasarnya. Dan untuk perbedaan dari proses kerja *stemmer* yang dibangun adalah pada *stemmer* tertentu memiliki proses tambahan seperti pengecekan kamus, *recoding,* dan pengembalian akhiran yang mensyaratkan proses tersebut pada algoritmanya.

Untuk proses pengecekan pada kamus dilakukan dengan metode *Binary Search (iterative)*. Langkah-langkah dari *binary search* mengikuti aturan dari *pseudocode* pada Gambar 4.3. Kata yang diproses akan dibandingkan dengan kata pada file teks kamus kata dasar, dimana proses pembandingannya diawali dengan *transformasi* data kata pada file teks kamus kata dasar kedalam bentuk *list* dan selanjutnya akan dicek kata yang diuji satu persatu kesesuaiannya dengan kata dalam *list* kamus. Apabila kata uji ditemukan pada kamus maka akan dikeluarkan sebagai kata dasarnya. Jika tidak, maka kata akan kembali diproses dengan langkah *stemming* selanjutnya maupun aturan pemenggalan yang lainnya.

|  |
| --- |
| 1. **function** BinarySearch(list[], min, max, key): 2. **while** max ≥ min **do** 3. mid = (min + max)/2 4. **if** list[mid] == key **then** 5. **return** mid 6. **else** **if** list[mid] **>** key **then** 7. mid = mid-1 8. **else** 9. mid = mid+1 10. **end** **if** 11. **end** **while** 12. **return** **False** |

Gambar 4.3 *Pseudocode Binary Search (iterative)*

# IMPLEMENTASI

## Lingkungan Implementasi

Implementasi dari sistem ini menggunakan bahasa pemrograman Python. Adapun peralatan dan bahan yang digunakan dalam implementasi ini adalah sebagai berikut:

1. Processor : AMD A8-6410 APU 2.00 GHz
2. RAM : 8 GB
3. Storage : SSD 240GB

Spesifikasi perangkat lunak yang digunakan untuk proses implementasi *stemmer* Bahasa Kawi menggunakan algoritma *Enhanced Confix Stripping* dan *Porter* antara lain.

1. Sistem Operasi : Windows 10 64-bit
2. Bahasa Pemrograman : Python 3.6.0
3. IDE : Spyder

## Implementasi Pengumpulan Data

Pengumpulan data yang dilakukan terbagi menjadi 3 kelompok dengan metode yang berbeda pada masing-masing kelompok. Berdasarkan subbab 4.3, data tersebut antara lain data kamus, data uji, dan data kata kunci.

### Pengumpulan Data Kamus

Pengumpulan data dilakukan dengan metode *scraping* seperti yang sudah dijelaskan pada anak subbab 4.3.1 pada situs *http://sealang.net/ojed/*. *Scraping* dilakukan dengan memperhatikan tag yang memuat kata dasar dengan bantuan library *BeautifulSoup*. Pada kasus ini, terdapat pada bagian tag html “a” dimana kata dasarnya sebagai teks pada link tersebut. Untuk memisahkan dengan tag “a” lainnya, maka ditambahkan *attribute* “title” pada tag “a” sebagai pembeda dengan tag lainnya. Implementasi untuk proses *scraping* menggunakan kode pemrograman *python* dapat dilihat pada Gambar 5.1.

|  |
| --- |
| **import** requests  **from** sys **import** argv  **from** bs4 **import** BeautifulSoup  script, huruf, hasil = argv  kamus = []  title = []  url = 'http://sealang.net/ojed/search.pl?service=entry&form= short&facet=head&query=' + huruf  r = requests.get(url)  p\_text = r.text  soup = BeautifulSoup(p\_text,'html.parser')  **for** link **in** soup.find\_all("a"):  kawi = link.text  title.append(link['title'])  kamus.append(kawi)  temp=[]  **for** i **in** range(len(title)):  **if** title[i] == "See full entry with citations, below. OJED ©1982 KITLV":  temp.append(kamus[i])  saveFile = open(hasil,'w', encoding='utf-8')  **for** item **in** temp:  saveFile.write("%s\n" % item) |

**Gambar 5.1 Kode Program Pengumpulan Data dengan *Scraping***

Proses eksekusi kode diatas dilakukan dengan memasukkan *input* argumen berupa huruf abjad dalam karakter Bahasa Kawi sehingga akan masuk ke variabel *huruf* dan argumen nama file *output* untuk menyimpan data hasil dari *scraping*. Data *scraping* disimpan ke dalam format berkas .txt dengan *encoding utf-8* karena data teks memiliki karakter huruf khusus. Proses diulangi dengan mengganti argumen huruf dari masing-masing kelompok huruf pada kamus yang terdapat pada situs tesebut.

### Pengumpulan Data Uji

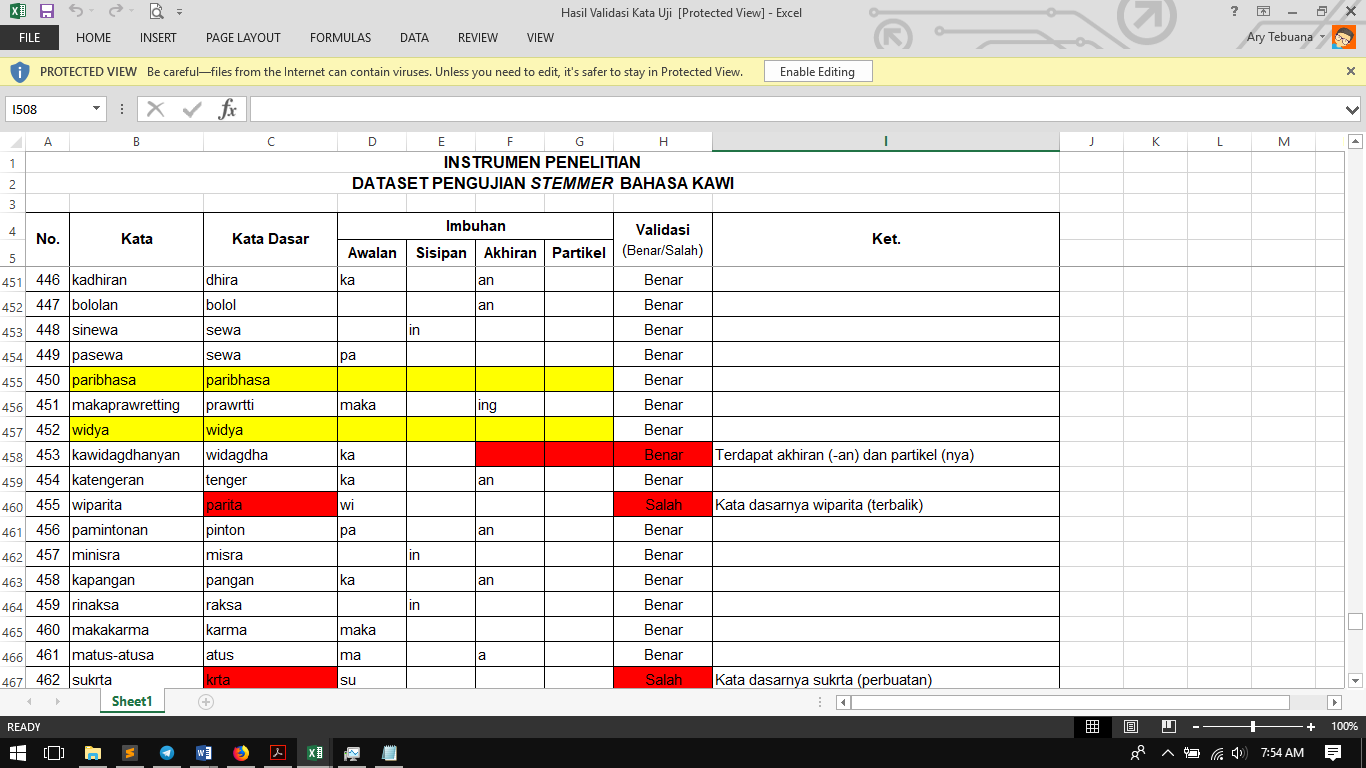
Proses pengumpulan data uji dilakukan secara manual dengan mendaftarkan seluruh kata pada sloka yang terdapat dalam kitab Sarasamuscaya. Format yang digunakan berbentuk huruf Latin yang disimpan dalam file teks (.txt). Untuk pengambilan kata uji dilakukan secara acak dengan memilih 430 kata berimbuhan dan 430 kata tanpa imbuhan. Dari proses pengambilan tersebut, seluruh kata yang terpilih dicocokan pola penulisan hurufnya ke dalam huruf Kawi. Cuplikan bentuk data uji dapat dilihat pada Gambar 5.2.

|  |
| --- |
| pinarāśrayan palungguhan tumiraskāra iwĕng santāpa dukut manghidĕp ahingsā mukti sipi waras dhyāna kawarjana manĕsĕl lumare mokṣa wĕkasan paṇḍita suraloka kaparārthan lepana śiṣṭācāra pĕngpönganikang sojar dhārmika kaśaktinika tuturan pinangguh dama rinambataning sĕkar nihan wyāpaka durśīla tapa kīrti tinapaknya tumingkah kapwan pakatoṅan sinyukan kṣetra śatru ĕngĕt |

Gambar 5.2 Cuplikan Data Uji

### Pengumpulan Data Kata Kunci

Data kata kunci dikumpulkan melalui pencocokkan data dari kata uji dengan kata yang terdapat pada Kamus Kawi – Bali. Analisis manual dilakukan dalam menentukan imbuhan yang digunakan apakah tergolong prefiks, infiks, sufiks, maupun partikel berdasarkan panduan aturan *wangun* dari sebuah kata. Data yang sudah dicocokkan akan dianalisis kembali oleh ahli yang telah disebutkan pada anak subbab 4.3.3, dimana hasil dari validasi kata kunci dari ahli memiliki bentuk seperti pada Gambar 4.1. Pada kolom validasi terdapat pernyataan Benar atau Salah yang diisi oleh ahli tersebut, jika terdapat kesalahan maka kesalahan tersebut akan dituliskan pada kolom keterangan beserta perbaikannya. Adapun contohnya seperti pada Gambar 5.3.



Gambar 5.3 Cuplikan Hasil Validasi Kata Kunci

## Implementasi Praproses Data

Tahap praproses data adalah tahap yang dilakukan sebelum dilakukan proses *stemming*. Dalam praproses terdapat 2 proses yang dilakukan, yaitu proses *case folding* dan tokenisasi.

### Implementasi *Case Folding*

Proses *case folding* merupakan tahapan untuk menormalisasikan bentuk huruf pada teks dalam dokumen. Untuk menghindari terjadinya kesalahan interpretasi pada arti kata yang diuji antara huruf kapital dengan huruf non-kapital yang berbeda, maka dilakukan penyeragaman bentuk huruf pada data teks menjadi huruf kecil (*lowercase*).

|  |
| --- |
| import re  def normalize\_text(text):  result = text.lower()  result = re.sub(r'[^āĕöīūḍṇṅñṛśṣṭa-z0-9 -]', ' ', result, flags = re.IGNORECASE|re.MULTILINE)  result = re.sub(r'( +)', ' ', result, flags = re.IGNORECASE|re.MULTILINE)  return result.strip() |

Gambar 5.4 Implementasi *Case Folding*

Pada Gambar 5.4 dapat dilihat penggunaan method *.lower()* untuk melakukan pengubahan bentuk huruf pada *input* teks menjadi huruf kecil. Selanjutnya teks yang tersimpan pada variabel *result* akan dicek dengan *regex* untuk memastikan tidak ada simbol lain selain yang bentuk alfabet yang diijinkan.

### Implementasi Tokenisasi

Pada proses tokenisasi dilakukan pemisahan terhadap kata-kata pembentuk kalimat pada teks yang dimasukkan. Kata-kata hasil normalisasi akan pisahkan menjadi token-token untuk selanjutnya dimasukkan pada *list*, sehingga dapat dilakukan proses *stemming* pada masing-masing kata.

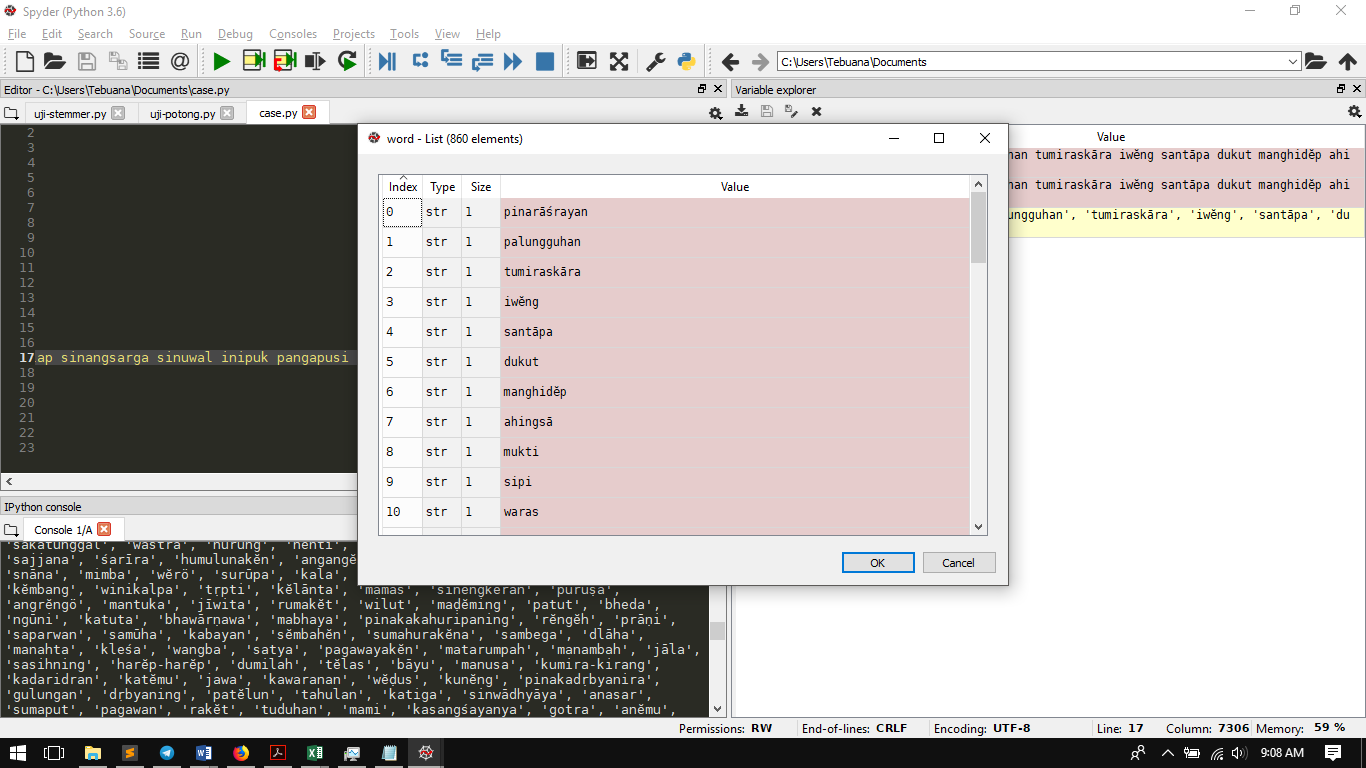
|  |
| --- |
| **def** stem(self, text):  normalizedText = TextNormalizer.normalize\_text(text)    words = normalizedText.split(' ')  stems = [] |

Gambar 5.5 Cuplikan Implementasi Tokenisasi (1)

|  |
| --- |
| **for** word **in** words:  stems.append(self.stem\_word(word))    **return** ' '.join(stems) |

Gambar 5.6 Cuplikan Implementasi Tokenisasi (2)

Pada baris ke-2 di Gambar 5.5, hasil dari normalisasi bentuk teks dengan *case folding* disimpan kedalam *variabel normalizedText*. Berikutnya pada baris ke-4 dilakukan pemisahan terhadap teks hasil normalisasi dengan melakukan method *.split()* pada *normalizedText* sehingga diperoleh kata yang terpisah satu dengan yang lainnya menjadi token-token. Proses ini bertujuan untuk membentuk token-token yang selanjutnya menjadi *list* pada *variabel* *word* seperti yang ditunjukan Gambar 5.7.



Gambar 5.7 Bentuk Data Kata Uji Setelah Tokenisasi

## Implementasi Algoritma *Enhanced Confix Stripping*

Penerapan dari algoritma *Enhanced Confix Stripping* pada *stemmer* Bahasa Kawi fokus implementasinya yaitu pada pengecekan kamus, penerapan aturan pemotongan imbuhan, dan method *loopPengembalianAkhiran*.Untuk *ECS* *stemmer* tanpa kamus akan terdapat perbedaan dimana tidak menggunakan bagian pengecekan kamus dan *loopPengembalianAkhiran*. Setiap kata yang diproses akan melalui pengecekan kata tunggal (*singular*) atau kata jamak (*plural*) terlebih dahulu. Kata yang jamak akan dipecah menjadi kata tunggal untuk di*stemming* dan jika memiliki kata dasar yang sama maka akan dikeluarkan sebagai hasil. Untuk implementasinya dapat dilihat pada Gambar 5.8.

|  |
| --- |
| **def** stem\_word(self, word):  **if** self.is\_plural(word):  **return** self.stem\_plural\_word(word)  **else**:  **return** self.stem\_singular\_word(word)    **def** is\_plural(self, word):  matches = re.match(r'^(.\*)-(ya|ang|ing|ni|ning|nikang|ikang |nta|nira|ika|nya|nyan)$', word)  **if** matches:  **return** matches.group(1).find('-') != -1  **return** word.find('-') != -1  **def** stem\_plural\_word(self, plural):  matches = re.match(r'^(.\*)-(.\*)$', plural)  **if** **not** matches:  **return** plural  words = [matches.group(1), matches.group(2)]  suffix = words[1]  suffixes = ['ya','ang','ing', 'ni', 'ning', 'nikang', 'ikang', 'nta', 'nira', 'ika', 'nya', 'nyan']  matches = re.match(r'^(.\*)-(.\*)$', words[0])  **if** suffix **in** suffixes **and** matches:  words[0] = matches.group(1)  words[1] = matches.group(2) + '-' + suffix  rootWord1 = self.stem\_singular\_word(words[0])  rootWord2 = self.stem\_singular\_word(words[1])  **if** rootWord1 == rootWord2:  **return** rootWord1  **else**:  **return** plural  **def** stem\_singular\_word(self, word):  context = Context(word, self.dictionary, self.visitor\_ provider)  context.execute()  **return** context.result |

Gambar 5.8 Kode Program Pengecekan *Singular* atau *Plural*

### Pengecekan Kamus

Proses pengecekan kamus didahului dengan melakukan *load* pada file kamus. *Load* file dilakukan melalui pemanggilan argumen saat proses dieksekusi program sehingga nama file yang menjadi input akan men-*trigger* variabel kamus pada method *get\_words\_from\_file()*. Method pada *class* *StemmerFactory* tersebut dijalankan untuk melakukan proses pembacaan dan pengambilan data kamus yang selanjutnya disimpan pada variabel *words*. Untuk kode program pengambilan file kamus dapat dilihat pada Gambar 5.9.

|  |
| --- |
| **def** get\_words\_from\_file(self):  kamus = argv[2]  current\_dir = os.path.dirname(os.path.realpath(\_\_file\_\_))  dictionaryFile = current\_dir + '/data/' + kamus  **if** **not** os.path.isfile(dictionaryFile):  **raise** RuntimeError('File Kamus Tidak Ditemukan')  dictionaryContent = ''  **with** open(dictionaryFile, 'r', encoding="utf8") **as** f:  dictionaryContent = f.read()  **return** dictionaryContent.split('\n') |

Gambar 5.9 Kode Program Pengambilan File Kamus

Data yang tersimpan pada variabel *words* selanjutnya digunakan sebagai *object* pada *class* ArrayDictionary yang disimpan pada variabel *dictionary*. Ketika terdapat proses dengan pemanggilan method pada variabel *dictionary*, maka yang dijalankan adalah method pada *class* ArrayDictionary dengan data yang diolah adalah data kamus dari *object*-nya yaitu *words* dan data pengujian pada variabel *word.* Pada class ini, fungsi pengecekan kata uji dengan kamus dilakukan dengan memanggil fungsi *binary\_search* pada method *contains*, dimana jika kata uji (*word*) tidak bernilai *None* dari hasil pemotongan imbuhan maka akan dipanggil fungsi *binary*\_*search* dengan parameter kata uji (*word*), kamus yang sudah tersimpan dalam list (*words*), panjang list kamus (*r*) dan nilai awal list kamus (*l*). Selanjutnya dicek oleh fungsi *binary*\_*search* yang akan mencari kata pada *index* dari list kamus dengan proses yang sesuai dengan langkah-langkah dari algoritma *Binary Search* pada subbab 4.2 yang nantinya akan mengeluarkan nilai *index* dari kata jika ditemukan dan dicek apakah data dari nilai *index* sama dengan kata yang dicari (word). Pengecekan terhadap kata dan kamus akan menghasilkan keluaran yang berupa nilai *True* atau *False.* Hasil tersebut nantinya akan me-*return* kata yang diproses jika ditemukan kata dasarnya. Implementasi dari pengecekan kata pada kamus dengan *Binary Search* dapat dilihat pada Gambar 5.10 berikut ini.

|  |
| --- |
| **class** ArrayDictionary(object):  **def** contains(self, word):  **if** word **is not** None:  l = 0  r = len(self.words)-1  i = self.binary\_search(self.words, l, r, word)  **if** self.words[i] == word:  **return** word  **def** binary\_search(self, words, l, r, word):  **while** r >= l:  mid = l + (r - l)//2  **if** self.words[mid] == word:  return mid  **elif** self.words[mid] > word:  r = mid-1  **else:**  l = mid+1  **return** False |

Gambar 5.10 Cuplikan Method Pengecekan Kata pada Kamus

# HASIL DAN PEMBAHASAN

## Pengujian Sistem dan Hasil

Pada penelitian ini dilakukan uji perfoma dari *stemmer* yang menggunakan algoritma *Enhanced Confix Stripping* dan Porter. Data uji yang digunakan berjumlah 860 kata yang pengumpulannya telah dijelaskan pada anak subbab 5.2.2. Proporsi data yang digunakan yaitu 50:50 dimana 430 kata berimbuhan dan 430 kata tanpa imbuhan. Sistem *stemmer* masing-masing dijalankan sebanyak 10 kali pengujian. Sistem dan berkas-berkas pengujian dijalankan dengan pemanggilan melalui argumen pada konsol pyhton. Setiap hasil *stemming* disimpan pada format berkas .txt dan dilakukan evaluasi pada hasil yang diperoleh sesuai dengan pengimplementasian rancangan evaluasi yang telah ditetapkan pada subbab 5.6.

## Hasil Praproses Data

Pada tahap ini dilakukan proses *case folding* dan tokenisasi pada data kata uji. Seperti yang ditunjukan pada Gambar 4.3, dimana data kata uji tersebut akan dinormalisasikan terlebih dahulu sebelum masuk ke sistem *stemmer*.

### Hasil *Case Folding*

Pada tahapan ini, data yang menjadi masukkan diubah format bentuk hurufnya menjadi huruf kecil (*lowercase*). Selain itu, simbol-simbol yang tidak diperlukan akan dihilangkan. Cuplikan hasil dari data kata uji yang sudah melalui *case folding* dapat dilihat pada Gambar 6.1.

|  |
| --- |
| pinarāśrayan palungguhan tumiraskāra iwĕng santāpa dukut manghidĕp ahingsā mukti sipi waras dhyāna kawarjana manĕsĕl lumare mokṣa wĕkasan paṇḍita suraloka kaparārthan lepana śiṣṭācāra pĕngpönganikang sojar dhārmika kaśaktinika tuturan pinangguh dama rinambataning sĕkar nihan wyāpaka durśīla tapa kīrti tinapaknya tumingkah kapwan pakatoṅan sinyukan kṣetra śatru ĕngĕt ginawenya baladewa paphala bhāwacakra mādhurya kalinganya sujanma kasakitan swargapada āditya kṣaya karakĕtan tinamban kapangan upādhyāya rajata asring makanimitta kalobhan mayuṣa ṣaḍbhāwa mokṣamārga alulut parĕkana pahiriban mangalahang mulahakĕning pukĕt anglugas śabda sukha sasiki makakarma mawuwuh patuwuh lunggat kataman punggung paḍĕm manguwuh upadeśa yowana wastu jĕmah kapawananing śīrṇa |

Gambar 6.1 Cuplikan Data Hasil Proses *Case Folding*

Dari data tersebut terlihat hanya karakter yang sesuai dengan aturan *regex* pada Gambar 5.4 yang akan dikeluarkan pada proses tersebut. Setiap kata sudah dalam bentuk huruf kecil dan tidak terdapat karakter-karakter simbol.

### Hasil Tokenisasi

Pada tahapan tokenisasi, data hasil dari proses case folding akan bentuk token-token. Setiap kata akan dipisahkan dengan tanda petik (‘…’) dan koma (,) sebagai pemisah antar token. Cuplikan data hasil dari proses tokenisasi dapat dilihat pada Gambar 6.2

|  |
| --- |
| [‘pinarāśrayan’, ‘palungguhan’, ‘tumiraskāra’, ‘iwĕng’, ‘santāpa’, ‘dukut’, ‘manghidĕp’, ‘ahingsā’, ‘mukti’, ‘sipi’, ‘waras’, ‘dhyāna’, ‘kawarjana’, ‘manĕsĕl’, ‘lumare’, ‘mokṣa’, ‘wĕkasan’, ‘paṇḍita’, ‘suraloka’, ‘kaparārthan’, ‘lepana’, ‘śiṣṭācāra’, ‘pĕngpönganikang’, ‘sojar’, ‘dhārmika’, ‘kaśaktinika’, ‘tuturan’, ‘pinangguh’, ‘dama’, ‘rinambataning’, ‘sĕkar’, ‘nihan’, ‘wyāpaka’, ‘durśīla’, ‘tapa’, ‘kīrti’, ‘tinapaknya’, ‘tumingkah’, ‘kapwan’, ‘pakatoṅan’, ‘sinyukan’, ‘kṣetra’, ‘śatru’, ‘ĕngĕt’, ‘ginawenya’, ‘baladewa’, ‘paphala’, ‘bhāwacakra’, ‘mādhurya’, ‘kalinganya’, ‘sujanma’, ‘kasakitan’, ‘swargapada’, ‘āditya’, ‘kṣaya’, ‘karakĕtan’, ‘tinamban’, ‘kapangan’, ‘upādhyāya’, ‘rajata’, ‘asring’, ‘makanimitta’, ‘kalobhan’, ‘mayuṣa’, ‘ṣaḍbhāwa’, ‘mokṣamārga’, ‘alulut’, ‘parĕkana’, ‘pahiriban’] |

Gambar 6.2 Hasil Proses Tokenisasi

Dari Gambar 6.2, selanjutnya data tersebut akan menjadi nilai dari *list* pada variabel *word*. Hal ini disebabkan dari *stemmer* yangmana prosesnya adalah mengolah kata per kata. Sehingga dengan ditokenisasi dapat memastikan bahwa data yang akan di*stemming* sudah dalam bentuk satuan kata.

## Hasil Proses *Stemming*

Hasil keluaran dari masing-masing *stemmer* memiliki bentuk keluaran yang sama yaitu berbentuk baris dan tersimpan dalam format berkas .txt. Untuk cuplikan dari hasil keluaran yang diperoleh dari sistem *stemmer* yang dibangun dapat dilihat pada Gambar 6.3.

|  |
| --- |
| parāśraya lungguh tiraskāra iwĕng santāpa dukut hidĕp ahingsā mukti sipi waras dhyāna warjana manĕsĕl lare mokṣa wĕkas paṇḍita suraloka parārtha lepana śiṣṭācāra pĕngpöng sojar dhārmika śakti tutur pangguh dama rambat sĕkar nihan wyāpaka durśīla tapa kīrti tapak tingkah kapwa pakatoṅan sinyukan kṣetra śatru ĕngĕt gawe baladewa phala bhāwacakra mādhurya ling sujanma sakit swargapada āditya kṣaya rakĕt tamba pang upādhyāya rajata |

Gambar 6.3 Cuplikan Hasil Keluaran dari *Stemmer*

Dari hasil keluaran pada Gambar 6.3, kata-kata tersebut akan ditransformasikan kedalam bentuk baris terpisah antara satu kata dengan yang lainnya dengan langkah yang dijabarkan pada anak subbab 5.6.2. Setelah data ditransformasikan, maka selanjutnya dapat diuji untuk mengevaluasi hasil akurasi yang diperoleh.

## Hasil Evaluasi Waktu Proses

Pada evaluasi pengujian waktu proses dari masing-masing *stemmer* yang dibangun, didapatkan hasil seperti yang ditunjukkan pada Tabel 6.1. Pada kolom rata-rata diperoleh melalui Persamaan 4.3.

Tabel 6.1 Hasil Evaluasi Waktu Proses *Stemmer*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Pengujian ke-** | **WAKTU PROSES (satuan detik)** | | | | | | |
| **A** | **B** | **C** | **D** | **E** | **F** | **G** |
| 1 | 0,7800 | 1,4161 | 1,3251 | 1,2791 | 0,0270 | 0,0290 | 0,0270 |
| 2 | 0,7780 | 1,3691 | 1,2841 | 1,2671 | 0,0289 | 0,0260 | 0,0269 |
| 3 | 0,8121 | 1,3561 | 1,2601 | 1,2711 | 0,0299 | 0,0269 | 0,0279 |
| 4 | 0,7651 | 1,3701 | 1,2911 | 1,2791 | 0,0280 | 0,0269 | 0,0260 |
| 5 | 0,7990 | 1,3781 | 1,2771 | 1,2801 | 0,0269 | 0,0259 | 0,0290 |
| 6 | 0,7580 | 1,3921 | 1,3191 | 1,2781 | 0,0260 | 0,0259 | 0,0279 |
| 7 | 0,7691 | 1,3871 | 1,2681 | 1,2681 | 0,0279 | 0,0270 | 0,0270 |
| 8 | 0,7560 | 1,4041 | 1,2951 | 1,2561 | 0,0290 | 0,0260 | 0,0270 |
| 9 | 0,7570 | 1,4761 | 1,2661 | 1,2741 | 0,0330 | 0,0260 | 0,0280 |
| 10 | 0,7921 | 1,4331 | 1,2661 | 1,2941 | 0,0259 | 0,0270 | 0,0269 |
| **Rata-rata** | **0,7767** | **1,3982** | **1,2852** | **1,2747** | **0,0283** | **0,0267** | **0,0274** |

**Keterangan**

A : *Enhanced Confix Stripping*

B : *Enhanced Confix Stripping* tanpa Kamus

C : *Enhanced Confix Stripping* tanpa Kamus (tanpa *suffix* a,i)

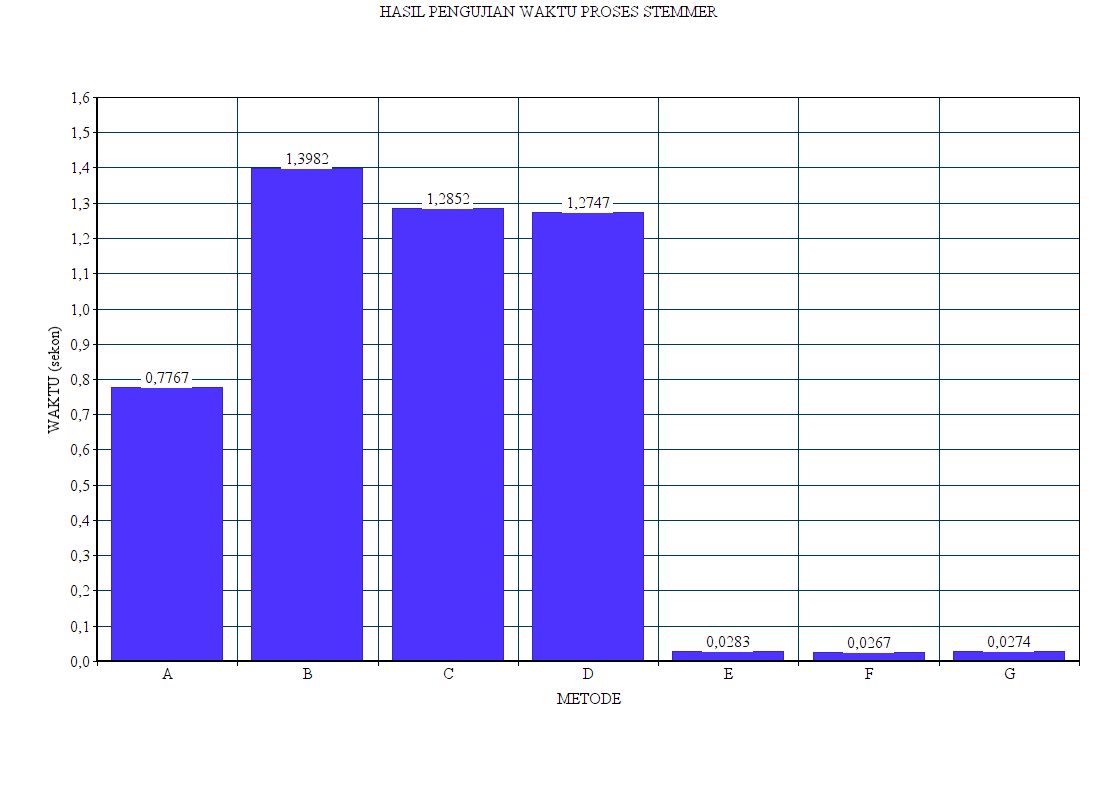
D : *Enhanced Confix Stripping* tanpa Kamus (tanpa *prefix* a, *suffix* a,i)

E : Porter

F : Porter (tanpa *suffix* a,i)

G : Porter (tanpa *prefix* a, *suffix* a,i)

Berdasarkan data hasil evaluasi terhadap waktu proses dari masing-masing metode yang diterapkan pada *Stemmer* Bahasa Kawi, maka dapat disajikan ke dalam bentuk grafik perbandingan rata-rata waktu proses dari *stemmer*. Berikut Gambar 6.4 adalah gambar grafik diagram batang yang merupakan interpretasi dari data Tabel 6.1, dimana masing-masing metode menggunakan keterangan yang sama pada tabel tersebut.



Gambar 6.4 Grafik Rata-rata Waktu Proses *Stemmer*

Pada Gambar 6.4 terlihat bahwa rata-rata waktu proses sistem yang dibutuhkan metode B, C, dan D paling lama dibandingkan metode lainnya yaitu dengan rata-rata waktu prosesnya 1,27 detik sampai dengan 1,39 detik, dimana metode B yaitu *Enhanced Confix Stripping* tanpa Kamus memiliki waktu proses terlama di angka 1,3982 detik. Hal tersebut dapat disebabkan oleh waktu yang dibutuhkan dari *stemmer* untuk menerapkan semua aturan pemenggalan pada kata yang diproses. Sedangkan, metode F yaitu Porter (tanpa *suffix* a,i) memiliki rata-rata waktu proses *stemming* paling cepat yaitu 0,0267 detik dimana setiap kata hanya dipotong imbuhannya saja jika ditemukan pada kata yang proses tanpa dilakukan pemeriksaan terhadap pola pembentukan dari kata tersebut.

Pada metode A yaitu *Enhanced Confix Stripping stemmer* relatif lebih lambat dibandingkan Porter *stemmer* yang menggunakan metode E, F, dan G dengan selisih rata-rata waktu proses *stemming-*nya ± 0,74 detik. Namun, metode A lebih cepat daripada metode B, C, dan D karena tidak semua aturan diterapkan pada kata uji yangmana ketika kata yang diproses sudah ditemukan di dalam kamus maka kata tersebut tidak dikenai aturan pemotongan imbuhan kembali dan proses *stemming* dihentikan.

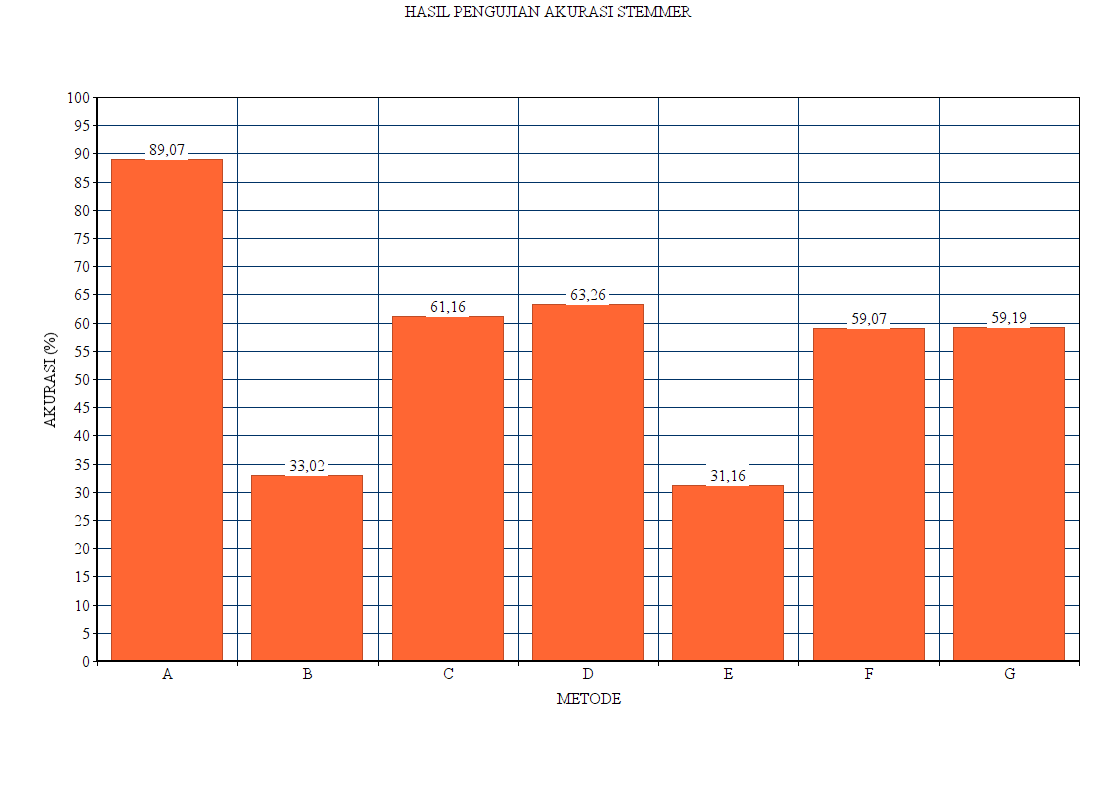
## Hasil Evaluasi Akurasi

Pada evaluasi akurasi, data keluaran dari masing-masing *stemmer* dilakukan pengecekan secara manual seperti yang sudah dijabarkan pada anak subbab 5.6.2. Dari hasil pengecekan pada kata kunci serta formulasi dengan persamaan yang digunakan maka diperoleh hasil seperti pada Tabel 6.2

Tabel 6.2 Hasil Evaluasi Akurasi *Stemmer*

|  |  |  |  |
| --- | --- | --- | --- |
| **Kode** | **Metode** | **Jumlah Kata Benar** | **Akurasi (%)** |
| A | *Enhanced Confix Stripping* | 766 | 89,07 |
| B | *Enhanced Confix Stripping* tanpa Kamus | 284 | 33,02 |
| C | *Enhanced Confix Stripping* tanpa Kamus (tanpa *suffix* a,i) | 526 | 61,16 |
| D | *Enhanced Confix Stripping* tanpa Kamus (tanpa *prefix* a, *suffix* a,i) | 544 | 63,26 |
| E | Porter | 268 | 31,16 |
| F | Porter (tanpa *suffix* a,i) | 508 | 59,07 |
| G | Porter (tanpa *prefix* a, *suffix* a,i) | 509 | 59,19 |

Berdasarkan Tabel 6.2, nilai jumlah kata benar adalah jumlah kata dari hasil masing-masing *stemmer* yang sesuai dengan kata kunci yang sudah divalidasi sebelumnya. Nilai akurasi dari masing-masing *stemmer* didapatkan dengan menggunakan perhitungan pada Persamaan 4.5 dan memasukkan nilai dari jumlah kata yang benar dari hasil *stemming* sehingga memperoleh persentase nilai akurasi tersebut. Berdasarkan Tabel 6.2, apabila data hasil evaluasi akurasi dari masing-masing *stemmer* diinterpretasikan ke dalam bentuk grafik diagram batang maka hasil yang didapatkan akan menjadi seperti pada Gambar 6.5, dimana untuk metode disimbolkan dengan huruf pada sesuai kode pada Tabel 6.2.



Gambar 6.5 Grafik Nilai Akurasi *Stemmer*

Pada Gambar 6.5 diperoleh bahwa tingkat akurasi tertinggi didapatkan pada metode A yaitu *Enhanced Confix Stripping* dengan jumlah kata yang sesuai kata kunci sebanyak 766 kata dari 860 kata uji dan akurasinya sebesar 89,07%. Sedangkan tingkat akurasi terendah diperoleh pada metode E yaitu Porter dengan jumlah kata yang benar sebanyak 268 kata dan akurasi sebesar 31.16%. Perbedaan yang cukup jauh disebabkan karena terjadinya *overstemming* pada metode yang tidak menggunakan kamus sehingga kata akan selalu dipotong hingga tidak ada lagi aturan pemotongan imbuhan yang memenuhi pada kata yang di-*stemming*.

Berdasarkan nilai akurasi yang didapatkan, terjadi peningkatan hasil sebesar ± 28% dengan penggunaan pengecualian terhadap beberapa aturan imbuhan untuk *Enhanced Confix Stripping stemmer* tanpa menggunakan kamus pada metode C dan D. Sedangkan, untuk Porter *stemmer* mendapatkan peningkatan nilai akurasi sebesar ± 27% jika menggunakan pengecualian aturan imbuhan seperti pada metode F dan G. Untuk hasil dari *stemmer* yang menggunakan pengecualian pada aturan pemotongan imbuhannya yaitu metode C, D, F, dan G sama-sama mendapatkan nilai akurasi diatas 59%, untuk *Enhanced Confix Stripping* tanpa Kamus dengan metode C dan D sedikit lebih unggul dibandingkan Porter dengan metode F dan G. Metode B (*Enhanced Confix Stripping* tanpa Kamus) dengan metode E (Porter)sama-sama memperoleh tingkat akurasi yang paling rendah. Untuk Porter *stemmer* pada metode E memiliki akurasi lebih rendah dibandingkan metode B dengan selisih 1.8% dimana nilai akurasi metode E yang diperoleh adalah sebesar 31.16%.

### Evaluasi Kesalahan *Stemming ECS*

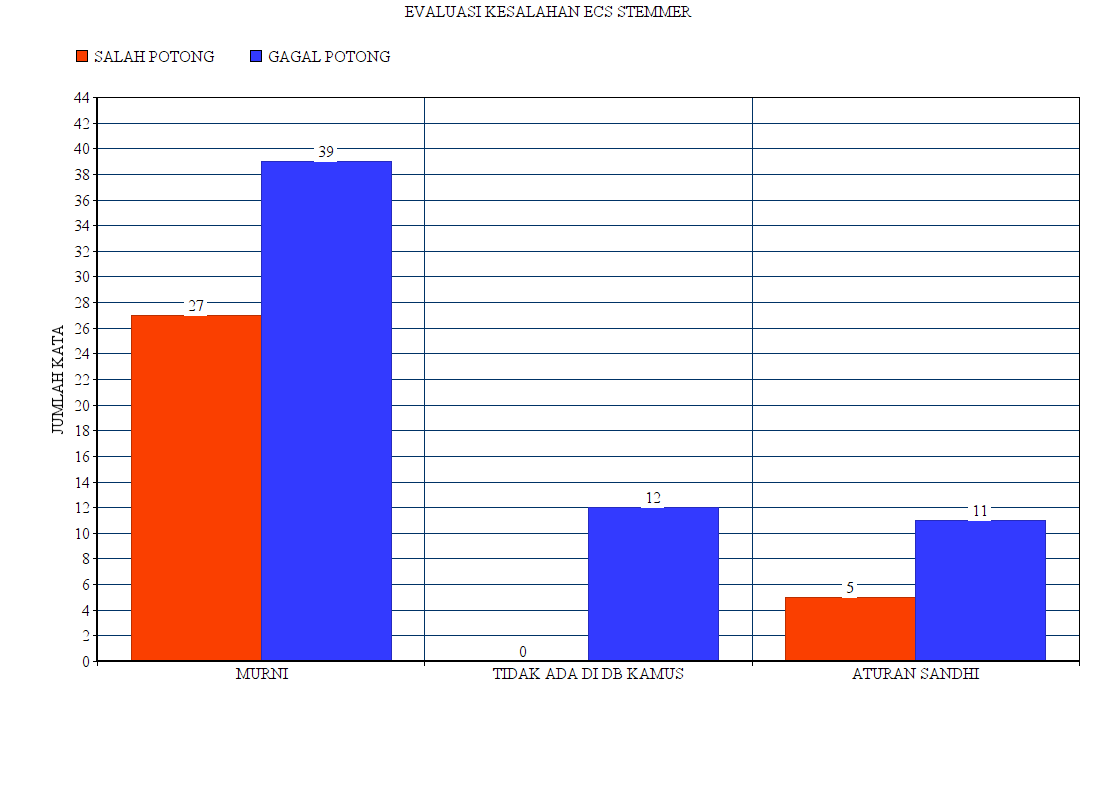
Proses evaluasi kesalahan pada hasil *stemming* dilakukan khusus untuk *Enhanced Confix Stripping* *stemmer* yang menggunakan kamus. Evaluasi dilakukan sesuai dengan kriteria-kriteria pada penjelasan anak subbab 5.6.2. Adapun hasil yang diperoleh dari evaluasi kesalahan *Enhanced Confix Stripping* *stemmer* untuk Bahasa Kawi seperti pada Tabel 6.3.

Tabel 6.3 Evaluasi Kesalahan Hasil *Stemming ECS*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Kategori I** | **Jumlah** | **Kategori II** | **Jumlah** | **Kategori III** | **Jumlah** |
| Salah Potong | 32 | Murni | 27 | *Overstemming* | 7 |
| *Understemming* | 4 |
| Lainnya | 16 |
| Tidak ada di kamus | 0 |  | |
| Aturan Sandhi | 5 |
| Gagal Potong | 62 | Murni | 39 | *Overstemming* | 0 |
| *Understemming* | 16 |
| Lainnya | 23 |
| Tidak ada di kamus | 12 |  | |
| Aturan Sandhi | 11 |

Berdasarkan Tabel 6.3, hasil *stemming* dengan kategori gagal potong memiliki jumlah yang lebih banyak dibandingkan kategori salah potong. Hal ini disebabkan oleh beberapa faktor. Untuk kategori gagal potong dipengaruhi oleh salah satunya yaitu aturan pemotongan yang tidak sesuai sehingga ketika proses pengecekan kamus selalu gagal mendapatkan kata dasar dan pada akhirnya kata dikembalikan ke bentuk awalnya sebagai *output*. Contohnya adalah kata “*dinohan*” yang seharusnya “*doh*” namun gagal diproses sehingga dikembalikan seperti bentuk semula. Sedangkan untuk kategori salah potong, salah satu faktor yang mempengaruhinya adalah ketika salah menerapkan aturan pemotongan namun kata ditemukan di kamus maka akan dikeluarkan sebagai *output*. Contohnya adalah kata “*kalaran*” yang seharusnya “*lara*” menjadi “*kalar*”.

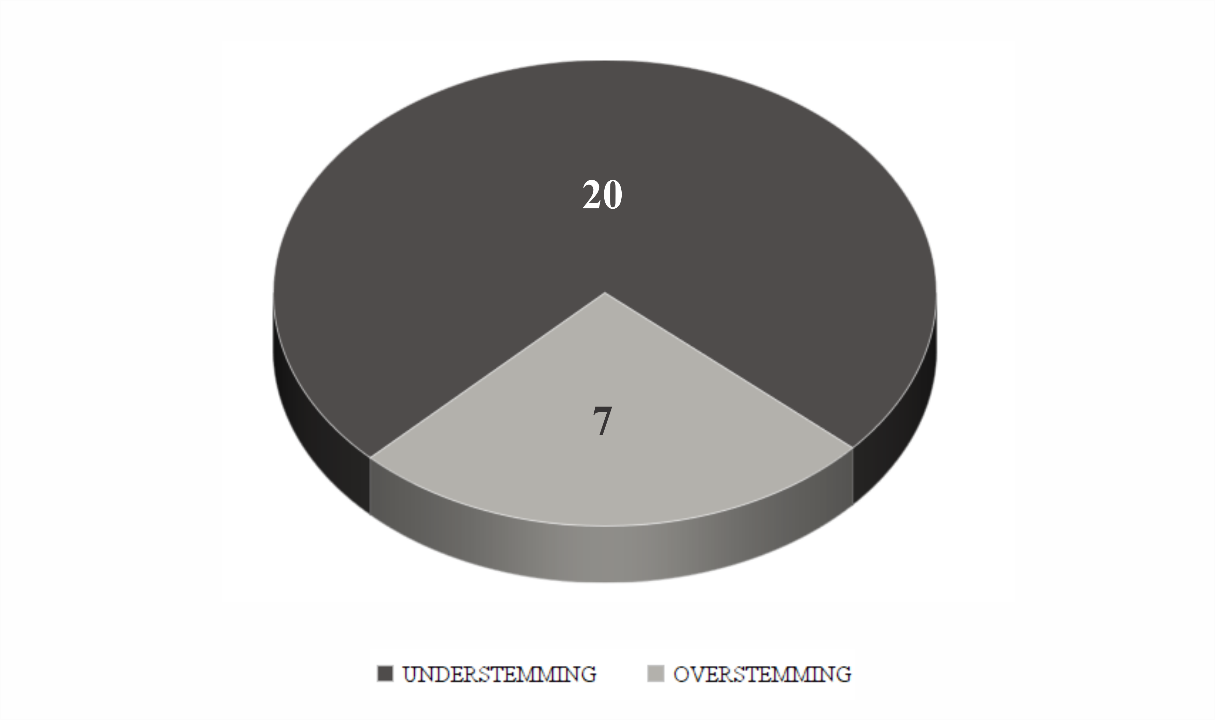
Masih mengenai masalah aturan pemotongan, hal yang mempengaruhi kesalahan pemotongan yaitu terdapat kata yang pembentukannya tergolong dalam aturan sandhi. Dalam *stemmer* yang dibangun tidak menggunakan aturan tersebut sehingga tidak ada aturan pemenggalan yang dapat melakukan pemotongan imbuhannya. Terdapat 5 kata untuk salah potong dan 11 kata untuk gagal potong. Contohnya “*karatun*” dengan kata dasar “*ratu*” yangmana huruf “*ĕ*” pada akhiran “-ĕn” melebur dengan huruf “u”. Selain itu, masalah ada tidaknya kata dasar pada data kamus yang digunakan berpengaruh terhadap hasil dari *stemming*nya, dimana untuk gagal potong terdapat 12 kata yang tidak ditemukan kata dasarnya di kamus. Dalam kasus ini tidak ditemukan pada kategori salah potong, dimana yang mungkin terjadi akibat ditemukannya kata yang panjang karakternya lebih pendek di kamus. Jika data evaluasi kesalahan *stemmer* pada Tabel 6.3 diinterpretasikan ke dalam sebuah grafik diagram batang maka akan terlihat seperti Gambar 6.6.



Gambar 6.6 Grafik Jumlah Kesalahan *Stemmer ECS*

Khusus pada hasil yang tergolong dalam kategori *salah potong* dan *murni*, dilakukan evaluasi kembali untuk menentukan jumlah kata yang tergolong kategori *overstemming* dan *understemming*. Dari data pada Tabel 6.3, terdapat 7 kata yang tergolong *overstemming*, dimana kata-kata tersebut mengalami pemotongan imbuhan yang berlebihan sehingga secara tidak disengaja kata dasarnya kembali dikenakan aturan pemotongan yang disebabkan karena hilangnya karakter huruf ketika pemotongan akhiran yang seharusnya tidak dipotong. Contohnya kata “*mangarjana*” menjadi “*arja*” yang seharusnya “*arjana*”, dimana sistem mendeteksi karakter “ana” sebagai akhiran yaitu “-an” dan “-a”. Untuk *understemming* terjadi ketika kata belum mencapai kata dasarnya tetapi sudah dikeluarkan sebagai kata dasar. Contohnya kata “*kasujanman*” yang seharusnya “*janma*” akan tetapi hasilnya menjadi “*sujanma*”. Pada kategori *lainnya* ditemukan sejumlah 16 kata yang menjadi masalah pada salah potong. Hal tersebut dikarenakan adanya *Word-sense disambiguation* yaitu satu kata yang dapat memiliki dua makna dan berasal dari kata dasar yang berbeda. Contohnya yaitu pada kata “*patĕkĕn*” bisa berarti “*tĕkĕn*” dan bisa juga “*patĕk*”, tergantung pada konteks kalimatnya. Namun, pada *stemmer* yang dibangun bersifat *context free grammar* sehingga tidak membatasi pada hasil produksi kata dari *stemmer*.

Selain pada kategori *salah potong* dan *murni*, kategori *gagal potong* dan *murni* juga dilakukan evaluasi kembali. Dari kata-kata yang tergolong *gagal potong* ditemukan masalah *understemming* sejumlah 16 kata. Hal ini disebabkan karena kata-kata tersebut sebelum dikenakan aturan pemotongan sudah ditemukan terlebih dahulu ketika proses pengecekan kata di kamus, sehingga kata mengalami gagal ­*stemming* dan dikeluarkan sebagai kata dasar. Untuk kategori *lainnya*, yang menjadi masalah yaitu ada tidaknya aturan pemotongan imbuhannya maupun kesesuaian kata dengan aturan yang pemotongan imbuhan yang dapat di-*stemming*. Contohnya yaitu pada kata “*madoh*” yang seharusnya “*doh*” karena tidak memenuhi syarat pada pemenggalan imbuhannya. Berikut pada Gambar 6.7 merupakan gambaran interpretasi dari hasil evaluasi terhadap kata yang tergolong dalam kategori *overstemming* dan *understemming* pada Tabel 6.3 secara keseluruhan tanpa memasukkan kategori *lainnya*.



Gambar 6.7 Diagram Evaluasi *Overstemming* dan *Understemming ECS*

Berdasarkan Gambar 6.7, diketahui bahwa hasil *stemming* yang tergolong *understemming* lebih dominan dibandingkan dengan *overstemming*. Jika diubah dalam bentuk persentase didapatkan perbandingan antara hasil *overstemming* dan *understemming* yaitu 7 kata berbanding 20 kata. Karena jumlah kata yang tergolong *understemming* yang lebih banyak dapat dikatakan bahwa *Enhanced Confix Stripping stemmer* yang dirancang untuk digunakan dalam Bahasa Kawi tergolong dalam jenis *light stemming* berdasarkan teori pada subbab 3.3.

### Evaluasi Kesalahan *Stemming ECS* tanpa Kamus dan Porter

Evaluasi kesalahan *stemming* pada metode *Enhanced Confix Stripping* tanpa Kamus dilakukan untuk mengetahui jumlah *overstemming* dan *understemming* yang dihasilkan. Seluruh metode yang menggunakan *Enhanced Confix Stripping* tanpa Kamus dan Porter dilakukan evaluasi, baik dengan maupun tanpa pengecualian aturan pemotongan imbuhan. Adapun hasil evaluasi yang didapatkan dapat dilihat pada Tabel 6.4.

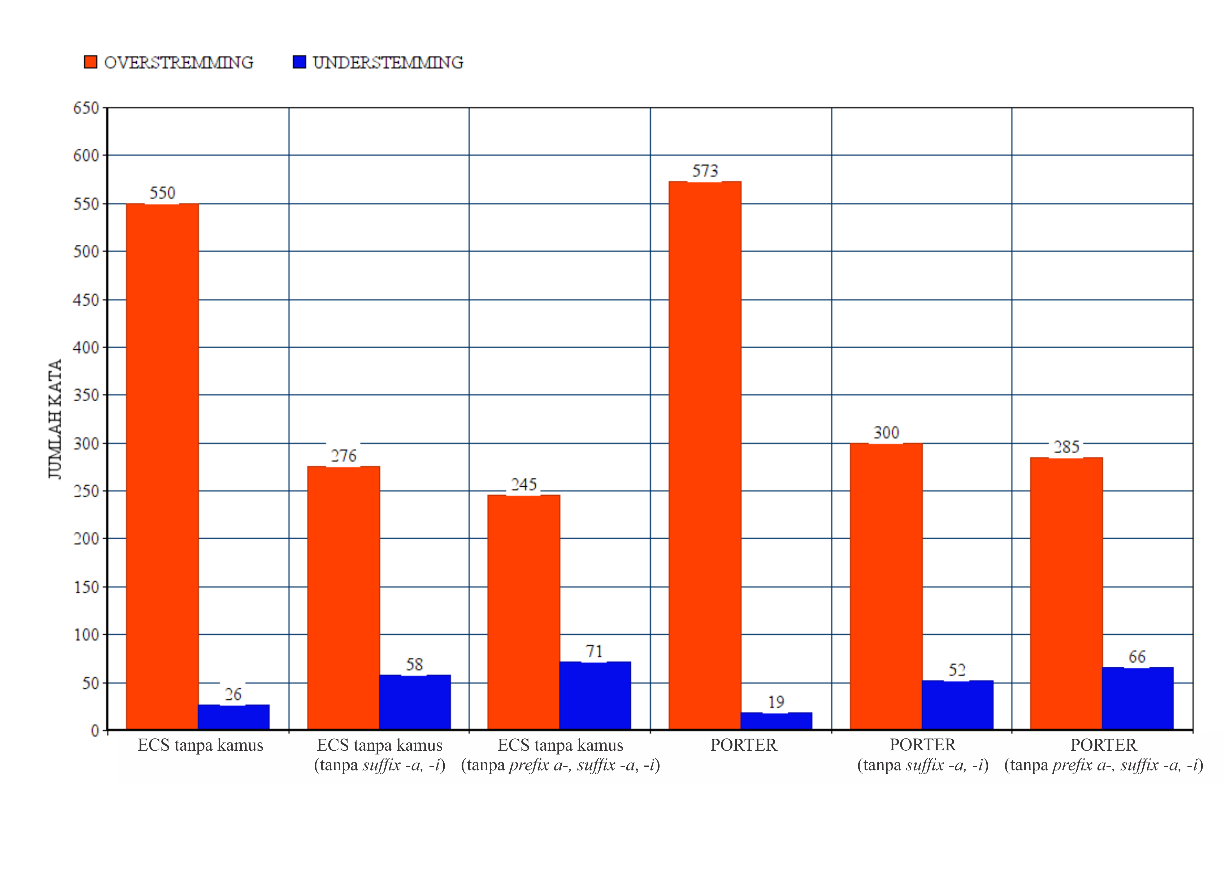
Tabel 6.4 Evaluasi Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter (1)

|  |  |  |
| --- | --- | --- |
| **Metode** | **Kesalahan Pemotongan** | |
| ***Overstemming*** | ***Understemming*** |
| *Enhanced Confix Stripping* tanpa Kamus | 550 | 26 |

Tabel 6.5 Evaluasi Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter (2)

|  |  |  |
| --- | --- | --- |
| **Metode** | **Kesalahan Pemotongan** | |
| ***Overstemming*** | ***Understemming*** |
| *Enhanced Confix Stripping* tanpa Kamus (tanpa *suffix* a,i) | 276 | 58 |
| *Enhanced Confix Stripping* tanpa Kamus (tanpa *prefix* a, *suffix* a,i) | 245 | 71 |
| Porter | 573 | 19 |
| Porter (tanpa *suffix* a,i) | 300 | 52 |
| Porter (tanpa *prefix* a, *suffix* a,i) | 285 | 66 |

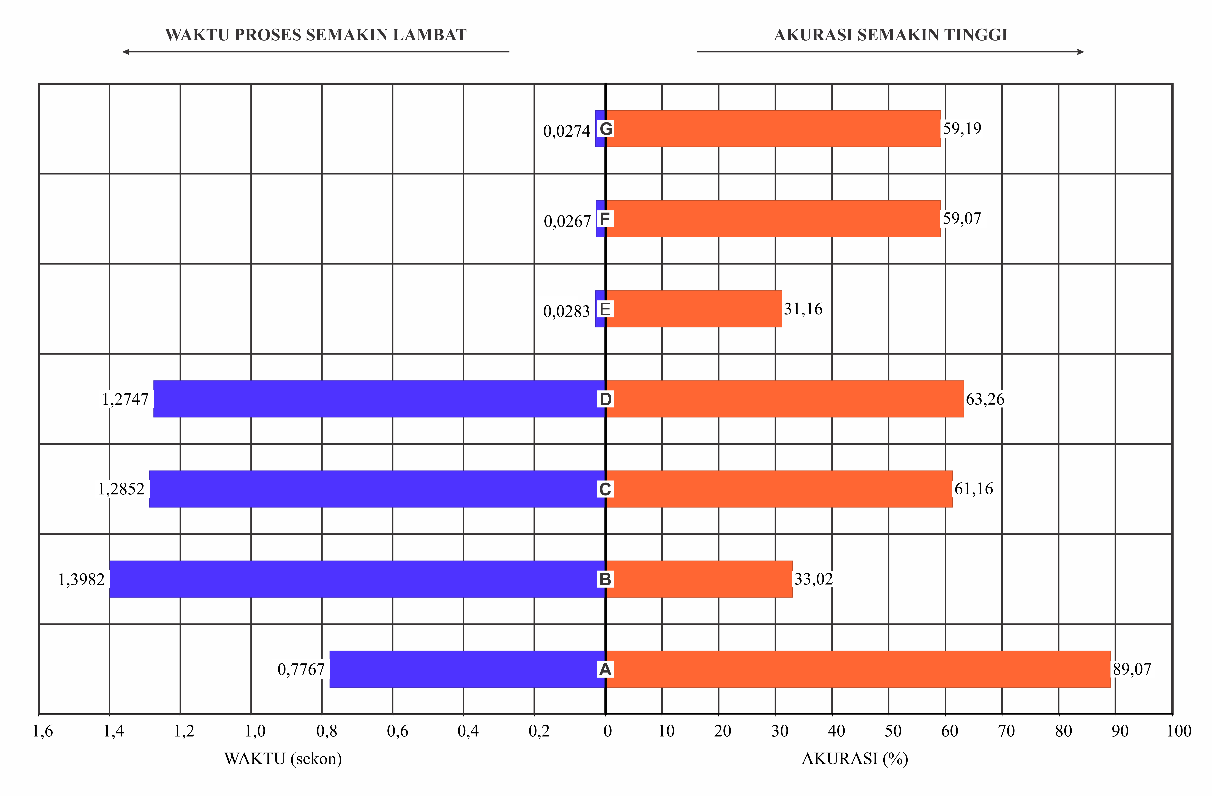
Berdasarkan hasil-hasil yang diperoleh pada Tabel 6.4, terlihat bahwa jumlah kata yang tergolong pada kategori *overstemming* lebih banyak daripada *understemming*. Apabila data pada Tabel 6.4 diinterpretasikan ke dalam bentuk grafik diagram batang maka akan menjadi seperti Gambar 6.8 dimana terlihat jumlah yang *overstemming* yang lebih banyak dibandingkan dengan *understemming* dari masing-masing metode yang digunakan. Dengan demikian dapat dikatakan metode *stemming* yang *Enhanced Confix Stripping* tanpa Kamus dan metode *stemming* Porter yang digunakan pada *stemmer* Bahasa Kawi termasuk dalam jenis *heavy stemming*.



Gambar 6.8 Diagram Evaluasi Tingkat Kesalahan *Stemming* *ECS* tanpa Kamus dan Porter

## Komparasi Hasil Evaluasi Waktu Proses dengan Akurasi

Hasil yang diperoleh dari masing-masing evaluasi *stemmer* yaitu waktu proses dan akurasi, selanjutnya dilakukan komparasi antara kedua hasil tersebut. Masing-masing metode akan dibandingkan hasil evaluasinya sehingga dapat dilihat performa dari metode *stemmer* yang digunakan untuk Bahasa Kawi. Berdasarkan data-data hasil evaluasi yang telah dilakukan, adapun hasil komparasi yang diperoleh dapat dilihat pada Gambar 6.9.



Gambar 6.9 Diagram Komparasi Hasil Evaluasi *Stemmer*

**Keterangan**

A : *Enhanced Confix Stripping*

B : *Enhanced Confix Stripping* tanpa Kamus

C : *Enhanced Confix Stripping* tanpa Kamus (tanpa *suffix* a,i)

D : *Enhanced Confix Stripping* tanpa Kamus (tanpa *prefix* a, *suffix* a,i)

E : Porter

F : Porter (tanpa *suffix* a,i)

G : Porter (tanpa *prefix* a, *suffix* a,i)

Pada Gambar 6.9, terlihat bahwa *stemmer* dengan menggunakan *Enhanced Confix Stripping* memiliki tingkat akurasi yang paling tinggi di angka 89.07% dengan rata-rata waktu proses yaitu 0,7767 detik. Sedangkan metode Porter *tanpa suffix a,i* dengan rata-rata waktu proses yang paling cepat yaitu 0,0267 detik memiliki tingkat akurasi angka 59,07%, dimana masih lebih rendah 2,09% dibandingkan dengan *Enhanced Confix Stripping* tanpa kamuspada metode yang sama yaitu *tanpa suffix a,i* namun dengan rata-rata waktu prosesnya yang lebih lama di angka 1,2852 detik. Sedangkan untuk *Enhanced Confix Stripping* tanpa Kamus memiliki rata-rata waktu proses terlama di angka 1,3982 detik dengan akurasi 33,02%. Akurasi paling rendah dimiliki oleh Porter dengan akurasi 31,16% dan rata-rata waktu proses 0,0283 detik.

Berdasarkan hasil komparasi antara evaluasi waktu proses dengan akurasi diperoleh bahwa dengan menggunakan metode pendekatan baik dengan atau tanpa kamus serta pengecualian pada aturan pemotongan imbuhan yang digunakan juga mempengaruhi waktu proses dan tingkat akurasi yang didapatkan dalam *stemming* Bahasa Kawi. Untuk *Enhanced Confix Stripping*, banyaknya jumlah aturan pemenggalan berpengaruh terhadap waktu proses, dimana penggunaan pengecekan kamus dapat membatasi proses penerapan aturan pada kata sehingga dapat memangkas waktu proses menjadi lebih cepat. Dan untuk *stemmer* tanpa kamus dengan pengecualian pemotongan imbuhan memiliki dampak pada tingkat akurasi yang diperoleh untuk Bahasa Kawi. Penggunaan *regular expression* pada *Enhanced Confix Stripping* berpengaruh pada waktu proses namun tingkat akurasi yang diperoleh lebih tinggi daripada *affix removal* pada Porter karena adanya bentuk pola yang harus dipenuhi sebelum pemotongan imbuhan dilakukan pada kata yang di-*stemming*.

# KESIMPULAN

## Kesimpulan

Dari penelitian mengenai metode *stemming* untuk Bahasa Kawi yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut:

1. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
2. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
3. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

## Saran

Berdasarkan penelitian yang telah dilakukan peneliti, terdapat beberapa saran yang dapat dikembangkan untuk penelitian selanjutnya antara lain.

1. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
2. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
3. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.
4. Aenean ultrices orci nec sem tempus, sed pretium erat ultricies. Proin quis nulla a magna mollis porta. Nullam in eleifend justo. Suspendisse non molestie dui, ac interdum risus.

#### DAFTAR PUSTAKA

Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S.M.M. dan Williams, H.E., 2007, Stemming Indonesian A Confix-Stripping Approach, *ACM Transactions on Asian Language Information Processing (TALIP)*, No. 3, Vol. 6, 1-33.

Agusta, L., 2009, Perbandingan Algoritma Stemming Porter Dengan Algoritma Nazief & Adriani Untuk Stemming Dokumen Teks Bahasa Indonesia, *Konferensi Nasional Sistem dan Informatika (KNS&I)*, Bali.

Alhanini, Y., dan Ab Aziz, M.J., 2011, The Enhancement of Arabic Stemming by Using Light Stemming and Dictionary-Based Stemming, *Journal of Software Engineering and Applications*, No. 9, Vol. 4, 522-526.

Arifin, A.Z., Mahendra, I.P.A.K., dan Ciptaningtyas, H.T., 2009, Enhanced Confix Stripping Stemmer and Ants Algorithm for Classifying News Document in Indonesian Language. *Proceeding of International Conference on Information & Communication Technology and Systems (ICTS)*, Melaka.

Fawcett, T., 2006, An Introduction to ROC Analysis, *Pattern Recognition Letters*, Vol. 27, 861-874.

Gitananda, W.A.S., 2017, Verba Nasalisasi N- (Atau Ng-?) Dalam Bahasa Bali Analisis Morfofonemik Dan Fungsi Sintaksis, *DHARMASMRTI,* No.1, Vol. XVI, 2-7.

Jivani, A.G., 2012, A Comparative Study of Stemming Algorithms, *International Journal of Control Theory and Applications*, No. 6, Vol. 2, 1930-1938

Kurniawan, B., Effendi, S. dan Sitompul O.S., 2012. Klasifikasi Konten Berita Dengan Metode Text Mining, *JURNAL DUNIA TEKNOLOGI INFORMASI*, No. 1, Vol. 1, 14-19.

Mardiwarsito, L., dan Kridalaksana, H., 2012, *Struktur Bahasa Jawa Kuna*, Edisi Kedua, Komunitas Bambu bekerja sama dengan École française d'Extrême-Orient, Depok.

Maulidi, R., 2016, Stemmer Untuk Bahasa Madura Dengan Modifikasi Metode Enhanced Confix Stripping Stemmer, *Prosiding Seminar Nasional Forum Dosen Indonesia*, Malang.

Miswanto, 2012, *Bahan Ajar Bahasa Kawi Untuk Mahasiswa STHD Klaten: MENGENAL BAHASA KAWI (1)*, STHD Klaten, Klaten.

Nata, G.N.M., dan Yudiastra, P.P., 2017, Stemming Teks Sor-Singgih Bahasa Bali, *Konferensi Nasional Sistem & Informatika*, Bali.

Purwoko, A., 2011, Model Stemming Berbasis Kamus Untuk Dokumen Berbahasa Sunda, *Tesis*, Sekolah Pascasarjana Institut Pertanian Bogor, Bogor.

Paice, C.D., 1996, Method for Evaluation of Stemming Algorithms Based on Error Counting, *Journal of the American Society for Information Science*, No. 8, Vol. 47, 632-649.

Ramanathan, A., dan Rao, D.D., 2003, A Lightweight Stemmer for Hindi, *the Proceedings of EACL*, Budapest.

Sharma, D., 2012, Stemming Algorithms A Comparative Study and their Analysis, *International Journal of Applied Information Systems (IJAIS)*, No. 3, Vol. 4, 7-12.

Sudana, O., Putra, D., Sudarma, M., Hartati, R.M. dan Diyastama, I.P.P., 2017, E-Translator Kawi to Balinese, *International Journal of Engineering and Emerging Technology*, No. 1, Vol. 2, 46-52.

Tahitoe, A.D. dan Purwitasari, D., 2010, Implementasi Modifikasi Enhanced Confix Stripping Stemmer Untuk Bahasa Indonesia Dengan Metode Corpus Based Stemming, *Skripsi*, Fakultas Teknologi Informasi Institut Teknologi Surabaya, Surabaya.

Tala, F.Z., 2003, A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia, *Tesis*, Institute for Logic, Language and Computation Universiteit van Amsterdam, Netherlands.

Teselkin, A.S., 1930, *Old Javanese (Kawi)* (diterjemahkan oleh Echols, J.M.), Penerbit Cornell Modern Indonesia Project, Ithaca.

Warna, I.W., dan Tim Penyusun, 1988, *Kamus Kawi – Bali*, Dinas Pendidikan Dasar Propinsi Daerah Tingkat I Bali, Denpasar.

Zoetmulder, P.J., 1994, *Kalangwan Sastra Jawa Kuno Selayang Pandang*, Djambatan, Jakarta.

Kuchling, A.M., 2016, Regular Expression HOWTO, *https://docs.python.org/release/3.6.0/howto/regex.html*, diakses tanggal 21 April 2018.

Indonesia, M., 2017, The Ancient Languages of Indonesia, *https://medium.com/@siwaratrikalpa/the-ancient-languages-of-indonesia-e44ef43579f5*, diakses tanggal 9 Oktober 2017.

#### LAMPIRAN

**LAMPIRAN 1**

**DATA UJI DAN KATA KUNCI BAHASA KAWI**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Kata** | **Kata Dasar** | **Imbuhan** | | | |
| **Prefiks** | **Infiks** | **Sufiks** | **Partkl.** |
| 1 | pinarāśrayan | parāśraya |  | in | an |  |
| 2 | palungguhan | lungguh | pa |  | an |  |
| 3 | tumiraskāra | tiraskāra |  | um |  |  |
| 4 | iwĕng | iwĕng |  |  |  |  |
| 5 | santāpa | santāpa |  |  |  |  |
| 6 | dukut | dukut |  |  |  |  |
| 7 | manghidĕp | hidĕp | ma |  |  |  |
| 8 | ahingsā | ahingsā |  |  |  |  |
| 9 | mukti | mukti |  |  |  |  |
| 10 | sipi | sipi |  |  |  |  |
| 11 | waras | waras |  |  |  |  |
| 12 | dhyāna | dhyāna |  |  |  |  |
| 13 | kawarjana | warjana | ka |  |  |  |
| 14 | manĕsĕl | sĕsĕl | ma |  |  |  |
| 15 | lumare | lare |  | um |  |  |
| 16 | mokṣa | mokṣa |  |  |  |  |
| 17 | wĕkasan | wĕkas |  |  | an |  |
| 18 | paṇḍita | paṇḍita |  |  |  |  |
| 19 | suraloka | suraloka |  |  |  |  |
| 20 | kaparārthan | parārtha | ka |  | an |  |
| 21 | lepana | lepana |  |  |  |  |
| 22 | śiṣṭācāra | śiṣṭācāra |  |  |  |  |
| 23 | pĕngpönganikang | pĕngpöng |  |  | an | ikang |
| 24 | sojar | ujar | sa |  |  |  |
| 25 | dhārmika | dhārmika |  |  |  |  |
| 26 | kaśaktinika | śakti | ka |  | ĕn | ika |
| 27 | tuturan | tutur |  |  | an |  |
| 28 | pinangguh | pangguh |  | in |  |  |
| 29 | dama | dama |  |  |  |  |
| 30 | rinambataning | rambat |  | in | an | ing |
| 31 | sĕkar | sĕkar |  |  |  |  |
| 32 | nihan | nihan |  |  |  |  |
| 33 | wyāpaka | wyāpaka |  |  |  |  |
| 34 | durśīla | durśīla |  |  |  |  |
| 35 | tapa | tapa |  |  |  |  |
| 36 | kīrti | kīrti |  |  |  |  |
| 37 | tinapaknya | tapak |  | in |  | nya |
| 38 | tumingkah | tingkah |  | um |  |  |
| 39 | kapwan | pwa | ka |  | an |  |
| 40 | pakatoṅan | katoṅ | pa |  | an |  |
| 41 | sinyukan | syuk |  | in | an |  |
| 42 | kṣetra | kṣetra |  |  |  |  |
| 43 | śatru | śatru |  |  |  |  |
| 44 | ĕngĕt | ĕngĕt |  |  |  |  |
| 45 | ginawenya | gawe |  | in |  | nya |
| 46 | baladewa | baladewa |  |  |  |  |
| 47 | paphala | phala | pa |  |  |  |
| 48 | bhāwacakra | bhāwacakra |  |  |  |  |
| 49 | mādhurya | mādhurya |  |  |  |  |
| 50 | kalinganya | eling |  |  |  |  |
| 51 | sujanma | janma | su |  |  |  |
| 52 | kasakitan | sakit | ka |  | an |  |
| 53 | swargapada | swargapada |  |  |  |  |
| 54 | āditya | āditya |  |  |  |  |
| 55 | kṣaya | kṣaya |  |  |  |  |
| 56 | karakĕtan | rakĕt | ka |  | an |  |
| 57 | tinamban | tamba |  | in | an |  |
| 58 | kapangan | pangan | ka |  | an |  |
| 59 | upādhyāya | upādhyāya |  |  |  |  |
| 60 | rajata | rajata |  |  |  |  |
| 61 | asring | asring |  |  |  |  |
| 62 | makanimitta | nimitta | maka |  |  |  |
| 63 | kalobhan | lobha | ka |  | an |  |
| 64 | mayuṣa | yuṣa | ma |  |  |  |
| 65 | ṣaḍbhāwa | ṣaḍbhāwa |  |  |  |  |
| 66 | mokṣamārga | mokṣamārga |  |  |  |  |
| 67 | alulut | lulut | a |  |  |  |
| 68 | parĕkana | parĕk |  |  | an+a |  |
| 69 | pahiriban | hirib | pa |  | an |  |
| 70 | mangalahang | kalah | ma |  | ang |  |
| 71 | mulahakĕning | ulah | ma |  | akĕn | ing |
| 72 | pukĕt | pukĕt |  |  |  |  |
| 73 | anglugas | lugas | a |  |  |  |
| 74 | śabda | śabda |  |  |  |  |
| 75 | sukha | sukha |  |  |  |  |
| 76 | sasiki | siki | sa |  |  |  |
| 77 | makakarma | karma | maka |  |  |  |
| 78 | mawuwuh | wuwuh | ma |  |  |  |
| 79 | patuwuh | tuwuh | pa |  |  |  |
| 80 | lunggat | lunggat |  |  |  |  |
| 81 | kataman | taman | ka |  |  |  |
| 82 | punggung | punggung |  |  |  |  |
| 83 | paḍĕm | paḍĕm |  |  |  |  |
| 84 | manguwuh | uwuh | ma |  |  |  |
| 85 | upadeśa | upadeśa |  |  |  |  |
| 86 | yowana | yowana |  |  |  |  |
| 87 | wastu | wastu |  |  |  |  |
| 88 | jĕmah | jĕmah |  |  |  |  |
| 89 | kapawananing | pawana | ka |  | an | ing |
| 90 | śīrṇa | śīrṇa |  |  |  |  |
| 91 | kulĕm | kulĕm |  |  |  |  |
| 92 | sapaku | aku | sa |  |  |  |
| 93 | pinungkurakĕnira | pungkur |  | in | akĕn | nira |
| 94 | hala | hala |  |  |  |  |
| 95 | āmbĕk | āmbĕk |  |  |  |  |
| 96 | paramārtha | paramārtha |  |  |  |  |
| 97 | iningĕti | ingĕt |  | in | i |  |
| 98 | parahu | parahu |  |  |  |  |
| 99 | wurung | wurung |  |  |  |  |
| 100 | surup | surup |  |  |  |  |
| 101 | kawidagdhanyan | widagdha | ka |  | an | nya |
| 102 | makapagwana | gwan | maka+pa |  | a |  |
| 103 | lĕnga | lĕnga |  |  |  |  |
| 104 | makasādhanang | sādhana | maka |  |  | ang |
| 105 | prang | prang |  |  |  |  |
| 106 | lālana | lālana |  |  |  |  |
| 107 | turū | turū |  |  |  |  |
| 108 | dĕmakan | dĕmak |  |  | an |  |
| 109 | kumuling | kumul |  |  |  | ing |
| 110 | paturwan | paturwan |  |  |  |  |
| 111 | ratna | ratna |  |  |  |  |
| 112 | mangabhiwāda | abhiwāda | ma |  |  |  |
| 113 | wacana | wacana |  |  |  |  |
| 114 | basa | basa |  |  |  |  |
| 115 | pasuruhan | suruh | pa |  | an |  |
| 116 | mitra | mitra |  |  |  |  |
| 117 | kewala | kewala |  |  |  |  |
| 118 | dharmaśāstra | dharmaśāstra |  |  |  |  |
| 119 | īrṣya | īrṣya |  |  |  |  |
| 120 | pāti | pāti |  |  |  |  |
| 121 | mada | mada |  |  |  |  |
| 122 | mala | mala |  |  |  |  |
| 123 | manuk | manuk |  |  |  |  |
| 124 | mangob | angob | ma |  |  |  |
| 125 | utpāta | utpāta |  |  |  |  |
| 126 | anglayang | layang | a |  |  |  |
| 127 | madoh | doh | ma |  |  |  |
| 128 | masādhana | sādhana | ma |  |  |  |
| 129 | pinujin | puji |  | in | en |  |
| 130 | kakĕnan | kĕna | ka |  | an |  |
| 131 | mapitutur | tutur | ma + pi |  |  |  |
| 132 | pamintonan | pinton | pa |  | an |  |
| 133 | niraya | niraya |  |  |  |  |
| 134 | yathāśakti | yathāśakti |  |  |  |  |
| 135 | kapūjān | pūjā | ka |  | an |  |
| 136 | matinggal | tinggal | ma |  |  |  |
| 137 | mangdadyakĕn | dadya | ma |  | akĕn |  |
| 138 | pūrwakarma | pūrwakarma |  |  |  |  |
| 139 | pawitra | pawitra |  |  |  |  |
| 140 | rubuh | rubuh |  |  |  |  |
| 141 | manunggang | tunggang | ma |  |  |  |
| 142 | kalasa | kalasa |  |  |  |  |
| 143 | haywa | haywa |  |  |  |  |
| 144 | mithya | mithya |  |  |  |  |
| 145 | pahuwusan | huwus | pa |  | an |  |
| 146 | hiniḍĕp | hiḍĕp |  | in |  |  |
| 147 | sulakṣaṇa | sulakṣaṇa |  |  |  |  |
| 148 | salah | salah |  |  |  |  |
| 149 | dudū | dudū |  |  |  |  |
| 150 | prabhu | prabhu |  |  |  |  |
| 151 | manggala | manggala |  |  |  |  |
| 152 | wĕnang | wĕnang |  |  |  |  |
| 153 | wiphala | phala | wi |  |  |  |
| 154 | pamalaku | palaku | pa |  |  |  |
| 155 | matuha | tuha | ma |  |  |  |
| 156 | durlabha | durlabha |  |  |  |  |
| 157 | tinutakĕn | tinut |  |  | akĕn |  |
| 158 | kabhuktyanya | bhukti | ka |  | an | nya |
| 159 | apuy | apuy |  |  |  |  |
| 160 | tumuntunakĕna | tuntun |  | um | akĕn+a |  |
| 161 | kinatṛṣṇān | tṛṣṇā | ka | in | an |  |
| 162 | lĕmbu | lĕmbu |  |  |  |  |
| 163 | pakarwa | rwa | paka |  |  |  |
| 164 | sādhu | sādhu |  |  |  |  |
| 165 | mānuṣa | mānuṣa |  |  |  |  |
| 166 | winawanya | wawa |  | in |  | nya |
| 167 | pañjut | pañjut |  |  |  |  |
| 168 | wicakṣaṇa | wicakṣaṇa |  |  |  |  |
| 169 | angusir | usir | a |  |  |  |
| 170 | pinakawāknira | wāk | pinaka |  |  | nira |
| 171 | pramāṇakĕna | pramāṇa |  |  | akĕn+a |  |
| 172 | tattwa | tattwa |  |  |  |  |
| 173 | nora | nora |  |  |  |  |
| 174 | abhra | abhra |  |  |  |  |
| 175 | masawah-sawah | sawah | ma |  |  |  |
| 176 | inupĕt | upĕt |  | in |  |  |
| 177 | sumur | sumur |  |  |  |  |
| 178 | swastha | swastha |  |  |  |  |
| 179 | duhkhamaya | duhkhamaya |  |  |  |  |
| 180 | pangatĕrakĕn | atĕr | pa |  | akĕn |  |
| 181 | angadĕg | adĕg | a |  |  |  |
| 182 | ijyā | ijyā |  |  |  |  |
| 183 | rumĕpa | rĕpa |  | um |  |  |
| 184 | madyusa | dyus | ma |  | a |  |
| 185 | pangrakṣa | rakṣa | pa |  |  |  |
| 186 | sĕkul | sĕkul |  |  |  |  |
| 187 | tumungkul | tungkul |  | um |  |  |
| 188 | kasambi | sambi | ka |  |  |  |
| 189 | kawandhana | wandha | ka |  | an+a |  |
| 190 | pangapahāsa | apahāsa | pa |  |  |  |
| 191 | sakawruhnya | wruh | sa+ka |  |  |  |
| 192 | sinuhan | suhan |  | in |  |  |
| 193 | tika | tika |  |  |  |  |
| 194 | sattwa | sattwa |  |  |  |  |
| 195 | niyata | niyata |  |  |  |  |
| 196 | inupakāra | upakāra |  | in |  |  |
| 197 | kimburu | kimburu |  |  |  |  |
| 198 | hĕlĕm | hĕlĕm |  |  |  |  |
| 199 | manah | manah |  |  |  |  |
| 200 | hiningān | hingā |  | in | an |  |
| 201 | patapa | tapa | pa |  |  |  |
| 202 | kabeh | kabeh |  |  |  |  |
| 203 | wwang | wwang |  |  |  |  |
| 204 | pāduka | pāduka |  |  |  |  |
| 205 | japa | japa |  |  |  |  |
| 206 | kawulatanika | wulat | ka |  | an | ika |
| 207 | humana | hana |  | um |  |  |
| 208 | wĕngi | wĕngi |  |  |  |  |
| 209 | sakaya-kaya | kaya | sa |  |  |  |
| 210 | pūrṇa | pūrṇa |  |  |  |  |
| 211 | kasantoṣan | santoṣa | ka |  | an |  |
| 212 | sukhaduhkha | sukhaduhkha |  |  |  |  |
| 213 | sinangguh | sangguh |  | in |  |  |
| 214 | patani | patani |  |  |  |  |
| 215 | mapuṇya | puṇya | ma |  |  |  |
| 216 | umĕt | umĕt |  |  |  |  |
| 217 | iṣṭasādhya | iṣṭasādhya |  |  |  |  |
| 218 | glāna | glāna |  |  |  |  |
| 219 | kamĕna | mĕna | ka |  |  |  |
| 220 | karĕsrĕs | rĕsrĕs | ka |  |  |  |
| 221 | kināryan | kārya |  | in | an |  |
| 222 | widyā | widyā |  |  |  |  |
| 223 | śiṣya | śiṣya |  |  |  |  |
| 224 | tumasakakĕning | tasak |  | um | akĕn | ing |
| 225 | tambā | tambā |  |  |  |  |
| 226 | manohara | manohara |  |  |  |  |
| 227 | lambā-lambān | lambā |  |  | an |  |
| 228 | pakawarĕgan | warĕg | paka |  | an |  |
| 229 | pramāda | pramāda |  |  |  |  |
| 230 | artha | artha |  |  |  |  |
| 231 | rinakṣa | rakṣa |  | in |  |  |
| 232 | sandhyā | sandhyā |  |  |  |  |
| 233 | sugandha | gandha | su |  |  |  |
| 234 | āyuṣa | āyuṣa |  |  |  |  |
| 235 | āgama | āgama |  |  |  |  |
| 236 | līna | līna |  |  |  |  |
| 237 | wiṣa | wiṣa |  |  |  |  |
| 238 | pĕtĕng | pĕtĕng |  |  |  |  |
| 239 | lĕbuh | lĕbuh |  |  |  |  |
| 240 | rākṣasa | rākṣasa |  |  |  |  |
| 241 | panglipura | lipur | pa |  | a |  |
| 242 | kinaragan | raga | ka | in | an |  |
| 243 | kadīrghāyuṣan | dīrghāyuṣa | ka |  | an |  |
| 244 | madya | madya |  |  |  |  |
| 245 | umĕgat | pĕgat |  | um |  |  |
| 246 | kinalakaran | lakar | ka | in | an |  |
| 247 | usungan | usung |  |  | an |  |
| 248 | rumĕgĕpang | rĕgĕp |  | um |  | ang |
| 249 | matuṇḍung | tuṇḍung | ma |  |  |  |
| 250 | mangusir | usir | ma |  |  |  |
| 251 | wuta | wuta |  |  |  |  |
| 252 | lĕnggang | lĕnggang |  |  |  |  |
| 253 | paturunan | turun | pa |  | an |  |
| 254 | bībhatsa | bībhatsa |  |  |  |  |
| 255 | śuddhaśīla | śuddhaśīla |  |  |  |  |
| 256 | bhojana | bhojana |  |  |  |  |
| 257 | pinakomah | umah | pinaka |  |  |  |
| 258 | purākṛta | purākṛta |  |  |  |  |
| 259 | umajarakĕn | ajar |  | um | akĕn |  |
| 260 | sawah | sawah |  |  |  |  |
| 261 | iwak | iwak |  |  |  |  |
| 262 | swarga | swarga |  |  |  |  |
| 263 | kapanasan | panas | ka |  | an |  |
| 264 | kayatnakĕna | yatna | ka |  | akĕn+a |  |
| 265 | timpalakĕna | timpal |  |  | akĕn+a |  |
| 266 | pawuni | wuni | pa |  |  |  |
| 267 | ayuṣa | yuṣa | a |  |  |  |
| 268 | kabrāhmaṇan | brāhmaṇa | ka |  | an |  |
| 269 | panastis | panastis |  |  |  |  |
| 270 | dharmasādhana | dharmasādhana |  |  |  |  |
| 271 | tandelan | tandel |  |  | an |  |
| 272 | tumamakĕn | tama |  | um | akĕn |  |
| 273 | pinarigrahanya | parigraha |  | in |  | nya |
| 274 | tumekakĕna | teka |  | um | akĕn+a |  |
| 275 | widhiwaśa | widhiwaśa |  |  |  |  |
| 276 | dadaha | dadaha |  |  |  |  |
| 277 | ārogya | ārogya |  |  |  |  |
| 278 | suluha | suluh |  |  | a |  |
| 279 | kalaran | lara | ka |  | an |  |
| 280 | tuli | tuli |  |  |  |  |
| 281 | śiṣṭa | śiṣṭa |  |  |  |  |
| 282 | nirbhāgya | bhāgya | nir |  |  |  |
| 283 | abhyudaya | abhyudaya |  |  |  |  |
| 284 | mamagĕhakĕn | pagĕh | ma |  | akĕn |  |
| 285 | inaku | aku |  | in |  |  |
| 286 | wṛddha | wṛddha |  |  |  |  |
| 287 | indriya | indriya |  |  |  |  |
| 288 | manungsung | sungsung | ma |  |  |  |
| 289 | inundang | undang |  | in |  |  |
| 290 | nṛśangsa | nṛśangsa |  |  |  |  |
| 291 | kasujanman | janma | ka | su | an |  |
| 292 | dinilatnya | dilat |  | in |  | nya |
| 293 | palĕman | alĕm | pa |  | an |  |
| 294 | drohaka | drohaka |  |  |  |  |
| 295 | sakatunggal | tunggal | sa+ka |  |  |  |
| 296 | wastra | wastra |  |  |  |  |
| 297 | hurung | hurung |  |  |  |  |
| 298 | hĕnti | hĕnti |  |  |  |  |
| 299 | rumĕngö | rĕngö |  | um |  |  |
| 300 | wipatha | wipatha |  |  |  |  |
| 301 | kapuhara | puhara | ka |  |  |  |
| 302 | sparśa | sparśa |  |  |  |  |
| 303 | sajjana | sajjana |  |  |  |  |
| 304 | śarīra | śarīra |  |  |  |  |
| 305 | humulunakĕn | hulun |  | um | akĕn |  |
| 306 | angangĕn-angĕn | angĕn | a |  |  |  |
| 307 | sasambhawa | sambhawa | sa |  |  |  |
| 308 | tinulungnya | tulung |  | in |  | nya |
| 309 | snāna | snāna |  |  |  |  |
| 310 | mimba | mimba |  |  |  |  |
| 311 | wĕrö | wĕrö |  |  |  |  |
| 312 | surūpa | rūpa | su |  |  |  |
| 313 | kala | kala |  |  |  |  |
| 314 | luṭik | luṭik |  |  |  |  |
| 315 | widyābhyāsa | widyābhyāsa |  |  |  |  |
| 316 | sep-sepan | sep |  |  | an |  |
| 317 | kĕmbang | kĕmbang |  |  |  |  |
| 318 | winikalpa | wikalpa |  | in |  |  |
| 319 | tṛpti | tṛpti |  |  |  |  |
| 320 | kĕlānta | kānta |  | el |  |  |
| 321 | mamās | mās | ma |  |  |  |
| 322 | sinĕngkĕran | sĕngkĕr |  | in | an |  |
| 323 | puruṣa | puruṣa |  |  |  |  |
| 324 | angrĕngö | rĕngö | a |  |  |  |
| 325 | mantuka | antuk | ma |  | a |  |
| 326 | jīwita | jīwita |  |  |  |  |
| 327 | rumakĕt | rakĕt |  | um |  |  |
| 328 | wilut | wilut |  |  |  |  |
| 329 | maḍĕming | paḍĕm | ma |  |  | ing |
| 330 | patut | patut |  |  |  |  |
| 331 | bheda | bheda |  |  |  |  |
| 332 | ngūni | ngūni |  |  |  |  |
| 333 | katuta | tuta | ka |  |  |  |
| 334 | bhawārṇawa | bhawārṇawa |  |  |  |  |
| 335 | mabhaya | bhaya | ma |  |  |  |
| 336 | pinakakahuripaning | hurip | pinaka+ka |  | an | ing |
| 337 | rĕngĕh | rĕngĕh |  |  |  |  |
| 338 | prāṇi | prāṇi |  |  |  |  |
| 339 | saparwan | parwa | sa |  | an |  |
| 340 | samūha | samūha |  |  |  |  |
| 341 | kabayan | kabayan |  |  |  |  |
| 342 | sĕmbahĕn | sĕmbah |  |  | ĕn |  |
| 343 | sumahurakĕna | sahur |  | um | akĕn+a |  |
| 344 | sambega | sambega |  |  |  |  |
| 345 | dlāha | dlāha |  |  |  |  |
| 346 | manahta | tahta | ma |  |  |  |
| 347 | kleśa | kleśa |  |  |  |  |
| 348 | wangba | wangba |  |  |  |  |
| 349 | satya | satya |  |  |  |  |
| 350 | pagawayakĕn | gawaya | pa |  | akĕn |  |
| 351 | matarumpah | tarumpah | ma |  |  |  |
| 352 | manambah | tambah | ma |  |  |  |
| 353 | jāla | jāla |  |  |  |  |
| 354 | sasihning | sasih |  |  |  | ning |
| 355 | harĕp-harĕp | harĕp |  |  |  |  |
| 356 | dumilah | dilah |  | um |  |  |
| 357 | tĕlas | tĕlas |  |  |  |  |
| 358 | bāyu | bāyu |  |  |  |  |
| 359 | manusa | manusa |  |  |  |  |
| 360 | kumira-kirang | kira |  | um |  | ang |
| 361 | kadaridran | daridra | ka |  | an |  |
| 362 | katĕmu | tĕmu | ka |  |  |  |
| 363 | jawa | jawa |  |  |  |  |
| 364 | kawaranan | waran | ka |  | an |  |
| 365 | wĕḍus | wĕḍus |  |  |  |  |
| 366 | kunĕng | kunĕng |  |  |  |  |
| 367 | pinakadṛbyanira | dṛbya | pinaka |  |  | nira |
| 368 | gulungan | gulung |  |  | an |  |
| 369 | dṛbyaning | dṛbya |  |  | an | ing |
| 370 | patĕlun | tĕlu | pa |  | ĕn |  |
| 371 | tahulan | tahulan |  |  |  |  |
| 372 | katiga | tiga | ka |  |  |  |
| 373 | sinwādhyāya | swādhyāya |  | in |  |  |
| 374 | anasar | nasar | a |  |  |  |
| 375 | sumaput | saput |  | um |  |  |
| 376 | pagawan | gawa | pa |  | an |  |
| 377 | rakĕt | rakĕt |  |  |  |  |
| 378 | tuduhan | tuduh |  |  | an |  |
| 379 | mami | mami |  |  |  |  |
| 380 | kasangśayanya | sangśaya | ka |  |  | nya |
| 381 | gotra | gotra |  |  |  |  |
| 382 | anĕmu | tĕmu | a |  |  |  |
| 383 | maling | maling |  |  |  |  |
| 384 | parumahaning | rumah | pa |  | an | ing |
| 385 | lobha | lobha |  |  |  |  |
| 386 | mamituhwa | mituhwa | ma |  |  |  |
| 387 | rahayu | rahayu |  |  |  |  |
| 388 | gandha | gandha |  |  |  |  |
| 389 | pagamĕlan | gamĕl | pa |  | an |  |
| 390 | wuwusakĕna | wuwus |  |  | akĕn+a |  |
| 391 | sinocan | soca |  | in | an |  |
| 392 | asambhawa | asambhawa |  |  |  |  |
| 393 | mṛṣāwāda | mṛṣāwāda |  |  |  |  |
| 394 | tunahan | tunah |  |  | an |  |
| 395 | alas | alas |  |  |  |  |
| 396 | duṣkṛta | duṣkṛta |  |  |  |  |
| 397 | dewatā | dewatā |  |  |  |  |
| 398 | wismṛti | wismṛti |  |  |  |  |
| 399 | kapāpan | pāpa | ka |  | an |  |
| 400 | rurū | rurū |  |  |  |  |
| 401 | sunde | sunde |  |  |  |  |
| 402 | kasiddha | siddha | ka |  |  |  |
| 403 | ākāśa | ākāśa |  |  |  |  |
| 404 | yaśa | yaśa |  |  |  |  |
| 405 | manĕhĕr | nĕhĕr | ma |  |  |  |
| 406 | kaparah | parah | ka |  |  |  |
| 407 | sulap | sulap |  |  |  |  |
| 408 | rinakwa | rakwa |  | in |  |  |
| 409 | rĕsĕp | rĕsĕp |  |  |  |  |
| 410 | mamara | mara | ma |  |  |  |
| 411 | kadweṣan | dweṣa | ka |  | an |  |
| 412 | inawaghāta | awaghāta |  | in |  |  |
| 413 | atyanta | atyanta |  |  |  |  |
| 414 | dinohan | doh |  | in | an |  |
| 415 | palagan | laga | pa |  | an |  |
| 416 | makambangan | kambang | ma |  | an |  |
| 417 | gṛhyastha | gṛhyastha |  |  |  |  |
| 418 | mamikul | mikul | ma |  |  |  |
| 419 | ikan | ikan |  |  |  |  |
| 420 | makolahang | kolah | ma |  |  | ang |
| 421 | swāmi | swāmi |  |  |  |  |
| 422 | paran | paran |  |  |  |  |
| 423 | dharma | dharma |  |  |  |  |
| 424 | umarahakĕn | arah |  | um | akĕn |  |
| 425 | kapāśa | pāśa | ka |  |  |  |
| 426 | kṣatriya | kṣatriya |  |  |  |  |
| 427 | hilang | hilang |  |  |  |  |
| 428 | pirak | pirak |  |  |  |  |
| 429 | inapit | apit |  | in |  |  |
| 430 | sangyoga | sangyoga |  |  |  |  |
| 431 | pamati-mati | mati | pa |  |  |  |
| 432 | kabañcana | bañcana | ka |  |  |  |
| 433 | kaprajñan | prajña | ka |  | an |  |
| 434 | weweh | weweh |  |  |  |  |
| 435 | rasa | rasa |  |  |  |  |
| 436 | wulangun | wulangun |  |  |  |  |
| 437 | lakṣmī | lakṣmī |  |  |  |  |
| 438 | kawangun | wangun | ka |  |  |  |
| 439 | purīṣa | purīṣa |  |  |  |  |
| 440 | agöng | agöng |  |  |  |  |
| 441 | kulawandhu | kulawandhu |  |  |  |  |
| 442 | agarawalan | garawal | a |  | an |  |
| 443 | tinunu | tunu |  | in |  |  |
| 444 | pagĕh | pagĕh |  |  |  |  |
| 445 | angidul | kidul | a |  |  |  |
| 446 | tamah | tamah |  |  |  |  |
| 447 | parama | parama |  |  |  |  |
| 448 | pahuntu | huntu | pa |  |  |  |
| 449 | ininggatan | inggat |  | in | an |  |
| 450 | langit | langit |  |  |  |  |
| 451 | anugraha | anugraha |  |  |  |  |
| 452 | karmaphala | karmaphala |  |  |  |  |
| 453 | matus-atusa | atus | ma |  | a |  |
| 454 | duśśīla | duśśīla |  |  |  |  |
| 455 | pamangguhanya | pangguh | pa |  | a | nya |
| 456 | ryak | ryak |  |  |  |  |
| 457 | kaparcaya | parcaya | ka |  |  |  |
| 458 | jĕnĕk | jĕnĕk |  |  |  |  |
| 459 | wukir | wukir |  |  |  |  |
| 460 | payang | payang |  |  |  |  |
| 461 | buddhi | buddhi |  |  |  |  |
| 462 | lalayan | lalay |  |  | an |  |
| 463 | saparĕk | parĕk | sa |  |  |  |
| 464 | bandhana | bandhana |  |  |  |  |
| 465 | sinakitan | sakit |  | in | an |  |
| 466 | tunggĕng | tunggĕng |  |  |  |  |
| 467 | luput | luput |  |  |  |  |
| 468 | tahĕn | tahĕn |  |  |  |  |
| 469 | mapupul | pupul | ma |  |  |  |
| 470 | masih | masih |  |  |  |  |
| 471 | nirmala | nirmala |  |  |  |  |
| 472 | manĕmbaha | sĕmbah | ma |  |  |  |
| 473 | sinambinya | sambi |  | in |  | nya |
| 474 | krodha | krodha |  |  |  |  |
| 475 | kṣamā | kṣamā |  |  |  |  |
| 476 | saduwĕg-duwĕga | duwĕg | sa |  | a |  |
| 477 | parĕng | parĕng |  |  |  |  |
| 478 | payung | payung |  |  |  |  |
| 479 | humaran-haran | haran |  | um |  |  |
| 480 | paganti | ganti | pa |  |  |  |
| 481 | bhinukti | bhukti |  | in |  |  |
| 482 | swāgata | swāgata |  |  |  |  |
| 483 | swangku | aku | swa+a |  |  |  |
| 484 | yukti | yukti |  |  |  |  |
| 485 | hurip | hurip |  |  |  |  |
| 486 | sininggahan | singgah |  | in | an |  |
| 487 | līlā | līlā |  |  |  |  |
| 488 | ātmā | ātmā |  |  |  |  |
| 489 | homa | homa |  |  |  |  |
| 490 | wwai | wwai |  |  |  |  |
| 491 | saguṇa | saguṇa |  |  |  |  |
| 492 | trikāya | trikāya |  |  |  |  |
| 493 | mamukti | mukti | ma |  |  |  |
| 494 | wyartha | wyartha |  |  |  |  |
| 495 | tĕmĕn | tĕmĕn |  |  |  |  |
| 496 | yajamāna | yajamāna |  |  |  |  |
| 497 | prihati | prihati |  |  |  |  |
| 498 | sukṛtakarma | sukṛtakarma |  |  |  |  |
| 499 | warga | warga |  |  |  |  |
| 500 | mamātya-mātyani | mātya | ma |  | i |  |
| 501 | rahina | rahina |  |  |  |  |
| 502 | eñjing | eñjing |  |  |  |  |
| 503 | madhura | madhura |  |  |  |  |
| 504 | tunggala | tunggal |  |  | a |  |
| 505 | kadang | kadang |  |  |  |  |
| 506 | brahmacāri | brahmacāri |  |  |  |  |
| 507 | kinawruhan | wruh | ka | in | an |  |
| 508 | kinĕnan | kĕna |  | in | an |  |
| 509 | upaśaman | upaśama |  |  | an |  |
| 510 | anitya | anitya |  |  |  |  |
| 511 | nirguṇa | nirguṇa |  |  |  |  |
| 512 | pakadangan | kadang | pa |  | an |  |
| 513 | tībra | tībra |  |  |  |  |
| 514 | hawan | hawan |  |  |  |  |
| 515 | abhyāsa | abhyāsa |  |  |  |  |
| 516 | mametwakĕna | metwa | ma |  | akĕn+a |  |
| 517 | nissangśaya | sangśaya | nir |  |  |  |
| 518 | padakṣiṇa | dakṣiṇa | pa |  |  |  |
| 519 | mangko | mangko |  |  |  |  |
| 520 | magawe | gawe | ma |  |  |  |
| 521 | upalakṣaṇa | lakṣaṇa | upa |  |  |  |
| 522 | mutya | mutya |  |  |  |  |
| 523 | lĕkasakĕna | lĕkas |  |  | akĕn+a |  |
| 524 | sinungkĕmaning | sungkĕm |  | in | an | ing |
| 525 | kumbha | kumbha |  |  |  |  |
| 526 | upawāsa | upawāsa |  |  |  |  |
| 527 | tribhuwana | tribhuwana |  |  |  |  |
| 528 | rajah | rajah |  |  |  |  |
| 529 | āyuṣya | āyuṣya |  |  |  |  |
| 530 | katekan | teka | ka |  | an |  |
| 531 | agirang | girang | a |  |  |  |
| 532 | kayogyaning | yogya | ka |  | an | ing |
| 533 | puruṣottama | puruṣottama |  |  |  |  |
| 534 | sarpa | sarpa |  |  |  |  |
| 535 | manastapa | nastapa | ma |  |  |  |
| 536 | bhagawān | bhagawān |  |  |  |  |
| 537 | amṛta | amṛta |  |  |  |  |
| 538 | krimi | krimi |  |  |  |  |
| 539 | giri-girinya | giri |  |  |  | nya |
| 540 | śāla | śāla |  |  |  |  |
| 541 | damĕlakĕnang | damĕl |  |  | akĕn | ang |
| 542 | pāruṣya | pāruṣya |  |  |  |  |
| 543 | kasatyan | satya | ka |  | an |  |
| 544 | sapinaranya | para | sa | in |  | nya |
| 545 | durgama | durgama |  |  |  |  |
| 546 | bhakta | bhakta |  |  |  |  |
| 547 | wibhawa | wibhawa |  |  |  |  |
| 548 | upastha | upastha |  |  |  |  |
| 549 | kagawayanya | gawaya | ka |  |  | nya |
| 550 | pinūjā | pūjā |  | in |  |  |
| 551 | magĕlĕm | gĕlĕm | ma |  |  |  |
| 552 | supātra | supātra |  |  |  |  |
| 553 | suwing | suwing |  |  |  |  |
| 554 | niṣkāraṇa | niṣkāraṇa |  |  |  |  |
| 555 | sulābhaning | lābha | su |  | an | ing |
| 556 | padulur | dulur | pa |  |  |  |
| 557 | āścarya | āścarya |  |  |  |  |
| 558 | prakāra | prakāra |  |  |  |  |
| 559 | wĕlĕkang | wĕlĕkang |  |  |  |  |
| 560 | prāṇa | prāṇa |  |  |  |  |
| 561 | inaranan | aran |  | in | an |  |
| 562 | abhijana | abhijana |  |  |  |  |
| 563 | maminta | pinta | ma |  |  |  |
| 564 | pĕngpĕngĕn | pĕngpĕng |  |  | ĕn |  |
| 565 | uminum | inum |  | um |  |  |
| 566 | makabratang | brata | maka |  | ang |  |
| 567 | makamusuh | musuh | maka |  |  |  |
| 568 | asungsung | sungsung | a |  |  |  |
| 569 | rūkṣa | rūkṣa |  |  |  |  |
| 570 | tinulak | tulak |  | in |  |  |
| 571 | tinĕmunya | tĕmu |  | in |  | nya |
| 572 | kalalah | lalah | ka |  |  |  |
| 573 | inucapan | ucap |  | in | an |  |
| 574 | prasada | prasada |  |  |  |  |
| 575 | palangka | palangka |  |  |  |  |
| 576 | gulinga | guling |  |  | a |  |
| 577 | makadoṣa | doṣa | maka |  |  |  |
| 578 | sāgara | sāgara |  |  |  |  |
| 579 | pasewa | sewa | pa |  |  |  |
| 580 | ilang | ilang |  |  |  |  |
| 581 | atĕngĕt | tĕngĕt | a |  |  |  |
| 582 | lawan | lawan |  |  |  |  |
| 583 | pinintan | pinta |  | in | an |  |
| 584 | kinalarakĕn | lara | ka | in | akĕn |  |
| 585 | inalĕm | alĕm |  | in |  |  |
| 586 | kĕḍĕr | kĕḍĕr |  |  |  |  |
| 587 | kinahanan | hana | ka | in | an |  |
| 588 | kāma | kāma |  |  |  |  |
| 589 | dumeh | dumeh |  |  |  |  |
| 590 | tĕkapan | tĕkap |  |  | an |  |
| 591 | sewaka | sewaka |  |  |  |  |
| 592 | ahangkāra | ahangkāra |  |  |  |  |
| 593 | mangwruhana | wruh | ma |  | an+a |  |
| 594 | lĕmĕh | lĕmĕh |  |  |  |  |
| 595 | panampak | tampak | pa |  |  |  |
| 596 | purih | purih |  |  |  |  |
| 597 | mataki-taki | taki | ma |  |  |  |
| 598 | rumuhun | ruhun |  | um |  |  |
| 599 | musuh | musuh |  |  |  |  |
| 600 | wiśeṣa | wiśeṣa |  |  |  |  |
| 601 | makapadārthang | padārtha | maka |  |  | ang |
| 602 | hilangakĕn | hilang |  |  | akĕn |  |
| 603 | pinakamĕng | amĕng | pinaka |  |  |  |
| 604 | muṇḍu | muṇḍu |  |  |  |  |
| 605 | pakṣa | pakṣa |  |  |  |  |
| 606 | akĕḍik | kĕḍik | a |  |  |  |
| 607 | winistarakĕn | wistara |  | in | akĕn |  |
| 608 | mārga | mārga |  |  |  |  |
| 609 | agĕlĕm | gĕlĕm | a |  |  |  |
| 610 | pūrṇama | pūrṇama |  |  |  |  |
| 611 | rumakṣa | rakṣa |  | um |  |  |
| 612 | malakang | malak |  |  |  | ang |
| 613 | wasana | wasana |  |  |  |  |
| 614 | mahuwusana | huwus | ma |  | an+a |  |
| 615 | kawarah | warah | ka |  |  |  |
| 616 | kakĕlĕm | kĕlĕm | ka |  |  |  |
| 617 | tinĕngĕtning | tĕngĕt |  | in |  | ning |
| 618 | inuttama | uttama |  | in |  |  |
| 619 | prastāwa | prastāwa |  |  |  |  |
| 620 | kṛtaghna | kṛtaghna |  |  |  |  |
| 621 | patah-patahning | patah |  |  |  | ning |
| 622 | lĕbu | lĕbu |  |  |  |  |
| 623 | kapĕtĕk | pĕtĕk | ka |  |  |  |
| 624 | sakta | sakta |  |  |  |  |
| 625 | gawe | gawe |  |  |  |  |
| 626 | dāna | dāna |  |  |  |  |
| 627 | kayowanan | yowana | ka |  | an |  |
| 628 | bala | bala |  |  |  |  |
| 629 | tasik | tasik |  |  |  |  |
| 630 | sakatambe | tambe | sa+ka |  |  |  |
| 631 | sākṣāt | sākṣāt |  |  |  |  |
| 632 | bhūmi | bhūmi |  |  |  |  |
| 633 | kaśmala | kaśmala |  |  |  |  |
| 634 | yayā | yayā |  |  |  |  |
| 635 | sangsāra | sangsāra |  |  |  |  |
| 636 | sumĕlat | sĕlat |  | um |  |  |
| 637 | durbuddhi | durbuddhi |  |  |  |  |
| 638 | desa | desa |  |  |  |  |
| 639 | durjana | durjana |  |  |  |  |
| 640 | mangarjana | arjana | ma |  |  |  |
| 641 | suteja | suteja |  |  |  |  |
| 642 | kaduṣṭan | duṣṭa | ka |  | an |  |
| 643 | wungkal | wungkal |  |  |  |  |
| 644 | maprayojana | prayojana | ma |  |  |  |
| 645 | triloka | triloka |  |  |  |  |
| 646 | kacuṇḍuka | cuṇḍuk | ka |  | a |  |
| 647 | nimeṣa | nimeṣa |  |  |  |  |
| 648 | pūrta | pūrta |  |  |  |  |
| 649 | hwat | hwat |  |  |  |  |
| 650 | tapwa | tapwa |  |  |  |  |
| 651 | tumulaka | tulak |  | um | a |  |
| 652 | makuru | kuru | ma |  |  |  |
| 653 | pinanganing | pangan |  | in |  | ing |
| 654 | pinanganya | pangan |  |  | an | nya |
| 655 | swāmī | swāmī |  |  |  |  |
| 656 | aiśwarya | aiśwarya |  |  |  |  |
| 657 | patamban | tamba | pa |  | an |  |
| 658 | patīrthan | tīrtha | pa |  | an |  |
| 659 | bololan | bololan |  |  |  |  |
| 660 | tumiba | tiba |  | um |  |  |
| 661 | paturū | turū | pa |  |  |  |
| 662 | śwadṛti | śwadṛti |  |  |  |  |
| 663 | strī | strī |  |  |  |  |
| 664 | ekaki | ekaki |  |  |  |  |
| 665 | alaris | laris | a |  |  |  |
| 666 | sinanga | sinang |  |  | a |  |
| 667 | kahīnan | hīna | ka |  | an |  |
| 668 | sukĕr | sukĕr |  |  |  |  |
| 669 | sukṛta | sukṛta |  |  |  |  |
| 670 | anātha | anātha |  |  |  |  |
| 671 | panĕngaha | tĕngah | pa |  | a |  |
| 672 | maṇik | maṇik |  |  |  |  |
| 673 | pamrihakĕna | pamrih |  |  | akĕn+a |  |
| 674 | angjanggut | janggut | a |  |  |  |
| 675 | niṣphala | niṣphala |  |  |  |  |
| 676 | iningūnira | ingū |  | in |  | nira |
| 677 | ahĕning | hĕning | a |  |  |  |
| 678 | kamokṣan | mokṣa | ka |  | an |  |
| 679 | kaparadāran | paradāra | ka |  | an |  |
| 680 | pāpakarma | pāpakarma |  |  |  |  |
| 681 | panĕmbĕlasanya | nĕmbĕlas | pa |  | an | nya |
| 682 | sinewa | sewa |  | in |  |  |
| 683 | marabi | rabi | ma |  |  |  |
| 684 | madhu | madhu |  |  |  |  |
| 685 | adharma | dharma | a |  |  |  |
| 686 | pinakolahnya | kolah | pa | in |  | nya |
| 687 | sinaputing | saput |  | in |  |  |
| 688 | pawilasan | wilasa | pa |  | an |  |
| 689 | agĕlis | gĕlis | a |  |  |  |
| 690 | niyama | niyama |  |  |  |  |
| 691 | herakĕn | her |  |  | akĕn |  |
| 692 | alara | lara | a |  |  |  |
| 693 | umambah | ambah |  | um |  |  |
| 694 | awasthā | awasthā |  |  |  |  |
| 695 | kaparwan | parwa | ka |  | an |  |
| 696 | yāma | yāma |  |  |  |  |
| 697 | kinatakutan | takut | ka | in | an |  |
| 698 | kagönganing | agöng | ka |  | an | ing |
| 699 | jñāna | jñāna |  |  |  |  |
| 700 | tumangguhi | tangguh |  | um | i |  |
| 701 | inahal | ahal |  | in |  |  |
| 702 | kawĕkas | wĕkas | ka |  |  |  |
| 703 | angucap | ucap | a |  |  |  |
| 704 | mantra | mantra |  |  |  |  |
| 705 | kāmuka | kāmuka |  |  |  |  |
| 706 | pĕpĕt | pĕpĕt |  |  |  |  |
| 707 | makawak | awak | maka |  |  |  |
| 708 | pangaji | aji | pa |  |  |  |
| 709 | nirwighna | wighna | nir |  |  |  |
| 710 | masambandha | sambandha |  |  |  |  |
| 711 | hibĕk | hibĕk |  |  |  |  |
| 712 | apanggih | panggih | a |  |  |  |
| 713 | saphala | phala | sa |  |  |  |
| 714 | mangisuhi | isuh | ma |  | i |  |
| 715 | makalayanang | layan | maka |  |  | ang |
| 716 | suputra | suputra |  |  |  |  |
| 717 | tarangga | tarangga |  |  |  |  |
| 718 | patĕkĕn | tĕkĕn | pa |  |  |  |
| 719 | kajĕnĕk | jĕnĕk | ka |  |  |  |
| 720 | jīrṇa | jīrṇa |  |  |  |  |
| 721 | kaduśśīlan | duśśīla | ka |  | an |  |
| 722 | kadyanggan | kadyangga |  |  | an |  |
| 723 | pangdūṣaṇa | dūṣaṇa | pa |  |  |  |
| 724 | kambulana | ambul | ka |  | an+a |  |
| 725 | durācāra | durācāra |  |  |  |  |
| 726 | mapagĕh | pagĕh | ma |  |  |  |
| 727 | tinanggap | tanggap |  | in |  |  |
| 728 | ārjawa | ārjawa |  |  |  |  |
| 729 | masamādhi | samādhi | ma |  |  |  |
| 730 | karikil | karikil |  |  |  |  |
| 731 | rājya | rājya |  |  |  |  |
| 732 | śarwa | śarwa |  |  |  |  |
| 733 | murungakĕn | murung |  |  | akĕn |  |
| 734 | rinĕngönta | rĕngö |  | in |  | nta |
| 735 | winehakĕnira | weh |  | in | akĕn | nira |
| 736 | majñāna | jñāna | ma |  |  |  |
| 737 | duṣṭa | duṣṭa |  |  |  |  |
| 738 | humilī | hilī |  | um |  |  |
| 739 | mārdawa | mārdawa |  |  |  |  |
| 740 | pakaphalang | phala | pa+ka |  | ang |  |
| 741 | hinanakĕn | hana |  | in | akĕn |  |
| 742 | wiwekī | wiwekī |  |  |  |  |
| 743 | kojaranya | ujar | ka |  |  | nya |
| 744 | manghilangakĕn | hilang | ma |  | akĕn |  |
| 745 | kanista | nista | ka |  |  |  |
| 746 | uddhata | uddhata |  |  |  |  |
| 747 | lumaku | laku |  | um |  |  |
| 748 | ujarakĕna | ujar |  |  | akĕn+a |  |
| 749 | kadhīran | dhīra | ka |  | an |  |
| 750 | wṛddhi | wṛddhi |  |  |  |  |
| 751 | kumĕtĕr | kĕtĕr |  | um |  |  |
| 752 | narakalokan | narakaloka |  |  | an |  |
| 753 | tingkah | tingkah |  |  |  |  |
| 754 | kinaniścayan | niścaya | ka | in | an |  |
| 755 | abhimāna | abhimāna |  |  |  |  |
| 756 | parcaya | parcaya |  |  |  |  |
| 757 | sabhā | sabhā |  |  |  |  |
| 758 | binabad | babad |  | in |  |  |
| 759 | manulung | tulung | ma |  |  |  |
| 760 | śaraṇa | śaraṇa |  |  |  |  |
| 761 | umah | umah |  |  |  |  |
| 762 | kulaja | kulaja |  |  |  |  |
| 763 | bhogopabhoga | bhogopabhoga |  |  |  |  |
| 764 | inārambha | ārambha |  | in |  |  |
| 765 | drawa | drawa |  |  |  |  |
| 766 | papa | papa |  |  |  |  |
| 767 | pinanasan | panas |  | in | an |  |
| 768 | winarah | warah |  | in |  |  |
| 769 | umantyakĕn | antya |  | um | akĕn |  |
| 770 | bĕsur | bĕsur |  |  |  |  |
| 771 | bhikṣuka | bhikṣuka |  |  |  |  |
| 772 | rare | rare |  |  |  |  |
| 773 | titikṣā | titikṣā |  |  |  |  |
| 774 | rumuh | rumuh |  |  |  |  |
| 775 | pahahĕning | hĕning | paha |  |  |  |
| 776 | maganti | ganti | ma |  |  |  |
| 777 | sugih | sugih |  |  |  |  |
| 778 | pĕkĕn | pĕkĕn |  |  |  |  |
| 779 | humĕnĕng | hĕnĕng |  | um |  |  |
| 780 | mata | mata |  |  |  |  |
| 781 | kinawaśakĕn | kawaśa |  | in | akĕn |  |
| 782 | gĕsĕng | gĕsĕng |  |  |  |  |
| 783 | paribhāṣa | paribhāṣa |  |  |  |  |
| 784 | subhikṣā | bhikṣā | su |  |  |  |
| 785 | panaḍahan | taḍah | pa |  | an |  |
| 786 | mahapuruṣa | puruṣa | maha |  |  |  |
| 787 | wijil | wijil |  |  |  |  |
| 788 | wiwāha | wiwāha |  |  |  |  |
| 789 | tulus | tulus |  |  |  |  |
| 790 | katiwasan | tiwas | ka |  | an |  |
| 791 | tamuy | tamuy |  |  |  |  |
| 792 | sasolahnya | solah | sa |  |  | nya |
| 793 | katĕngĕran | tĕngĕr | ka |  | an |  |
| 794 | maśarīra | śarīra | ma |  |  |  |
| 795 | umakusāra | akusāra |  | um |  |  |
| 796 | asing | asing |  |  |  |  |
| 797 | daridra | daridra |  |  |  |  |
| 798 | loka | loka |  |  |  |  |
| 799 | prakīrti | prakīrti |  |  |  |  |
| 800 | landĕping | landĕp |  |  | ing |  |
| 801 | unggu | unggu |  |  |  |  |
| 802 | kuṇḍa | kuṇḍa |  |  |  |  |
| 803 | śūdra | śūdra |  |  |  |  |
| 804 | pilih | pilih |  |  |  |  |
| 805 | mamṛtyu | mṛtyu | ma |  |  |  |
| 806 | kinĕlikaning | kĕlik |  | in | an | ing |
| 807 | hinayu | hayu |  | in |  |  |
| 808 | mangĕna | angĕn | ma |  | a |  |
| 809 | nirwikāra | nirwikāra |  |  |  |  |
| 810 | suminggahanang | singgah |  | um | an | ang |
| 811 | amĕngpönga | pĕngpöng | a |  | a |  |
| 812 | pangalap | alap | pa |  |  |  |
| 813 | tātan | tātan |  |  |  |  |
| 814 | wyālasarpa | wyālasarpa |  |  |  |  |
| 815 | linakṣaṇan | lakṣaṇa |  | in | an |  |
| 816 | kasuban | suba | ka |  | an |  |
| 817 | tīrthayātra | tīrthayātra |  |  |  |  |
| 818 | kasanggahanika | sanggah | ka |  | an | ika |
| 819 | irang | irang |  |  |  |  |
| 820 | tukar | tukar |  |  |  |  |
| 821 | sĕgĕh | sĕgĕh |  |  |  |  |
| 822 | tuminggalakĕn | tinggal |  | um | akĕn |  |
| 823 | amangan | mangan | a |  |  |  |
| 824 | matanghi | tanghi | ma |  |  |  |
| 825 | langgĕng | langgĕng |  |  |  |  |
| 826 | bhujālamba | bhujālamba |  |  |  |  |
| 827 | manglampuran | lampur | ma |  | an |  |
| 828 | tambak | tambak |  |  |  |  |
| 829 | kakawaca | kawaca | ka |  |  |  |
| 830 | rinampas | rampas |  | in |  |  |
| 831 | dodot | dodot |  |  |  |  |
| 832 | karatun | ratu | ka |  | en |  |
| 833 | samitra | mitra | sa |  |  |  |
| 834 | paramasukha | paramasukha |  |  |  |  |
| 835 | humiḍĕpang | hiḍĕp |  | um |  | ang |
| 836 | mūḍha | mūḍha |  |  |  |  |
| 837 | tuha | tuha |  |  |  |  |
| 838 | prajña | prajña |  |  |  |  |
| 839 | pajara | jara | pa |  |  |  |
| 840 | mokṣapada | mokṣapada |  |  |  |  |
| 841 | makaprawṛtting | prawṛtti | maka |  |  | ing |
| 842 | masukha-sukhan | sukha | ma |  | an |  |
| 843 | anggarāga | anggarāga |  |  |  |  |
| 844 | wulat | wulat |  |  |  |  |
| 845 | kĕlabakĕn | kĕlab |  |  | akĕn |  |
| 846 | inantining | anti |  | in | ĕn | ing |
| 847 | katakutnya | takut | ka |  |  | nya |
| 848 | wiparīta | wiparīta |  |  |  |  |
| 849 | nirwyādhi | wyādhi | nir |  |  |  |
| 850 | amuhara | muhara | a |  |  |  |
| 851 | inalap | alap |  | in |  |  |
| 852 | sinangsarga | sangsarga |  | in |  |  |
| 853 | sinuwal | suwal |  | in |  |  |
| 854 | inipuk | ipuk |  | in |  |  |
| 855 | pangapusi | apus | pa |  | i |  |
| 856 | tinaḍahnya | taḍah |  | in |  | nya |
| 857 | śrotriya | śrotriya |  |  |  |  |
| 858 | pātra | pātra |  |  |  |  |
| 859 | miniśra | miśra |  | in |  |  |
| 860 | angrĕpa | rĕpa | a |  |  |  |