

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
PERNYATAAN.....	ii
HALAMAN MOTTO DAN PERSEMBAHAN.....	iii
PRAKATA.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	x
INTISARI	xi
ABSTRACT	xii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Metode Penelitian.....	4
1.7. Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
BAB III LANDASAN TEORI.....	12
3.1. Bahasa Kawi.....	12
3.2. Teknik <i>Stemming</i>	21
3.2.1. Algoritma <i>Stemming</i> Nazief dan Adriani	23
3.2.2. Algoritma <i>Confix-Stripping Stemmer</i>	25
3.2.3. Algoritma <i>Enhanced Confix Stripping (ECS) Stemmer</i>	27
3.2.4. Algoritma Porter	31
3.3. Kesalahan <i>Stemming</i>	33
3.4. <i>Regular Expression</i>	34
BAB IV ANALISIS DAN PERANCANGAN.....	36
4.1. Analisis Permasalahan	36
4.2. Rancangan Umum Sistem.....	37
4.3. Data	39
4.3.1. Data Kamus.....	39
4.3.2. Data Kata Uji.....	40
4.3.3. Data Kata Dasar (Kunci Jawaban)	40
4.4. Praproses Data.....	41
4.4.1. <i>Case Folding</i>	41
4.4.2. Tokenisasi	42
4.5. Perancangan <i>Stemmer</i>	42
4.5.1. <i>Enhanced Confix Stripping Stemmer</i>	43
4.5.2. <i>Enhanced Confix Stripping Stemmer</i> Tanpa Kamus	50
4.5.3. Porter <i>Stemmer</i>	53
4.6. Perancangan Evaluasi.....	56

BAB V IMPLEMENTASI.....	61
5.1. Lingkungan Implementasi.....	61
5.2. Implementasi Pengumpulan Data	61
5.2.1. Pengumpulan Data Kamus	61
5.2.2. Pengumpulan Data Uji	62
5.2.3. Pengumpulan Data Kata Kunci	63
5.3. Implementasi Praproses Data	64
5.3.1. Implementasi <i>Case Folding</i>	64
5.3.2. Implementasi Tokenisasi.....	64
5.4. Implementasi Algoritma <i>Enhanced Confix Stripping</i>	65
5.4.1. Pengecekan Kamus	66
5.4.2. Penerapan Aturan Pemotongan Akhiran	68
5.4.3. Penerapan Aturan Pemotongan Awalan dan Sisipan	69
5.4.4. Penerapan <i>loopPengembalianAkhiran</i>	71
5.4.5. Penerapan Alur Proses Algoritma	73
5.5. Implementasi Algoritma Porter.....	74
5.6. Implementasi Pengujian dan Evaluasi	77
5.6.1. Evaluasi Waktu Proses	78
5.6.2. Evaluasi Akurasi	78
BAB VI HASIL DAN PEMBAHASAN	80
6.1. Pengujian Sistem dan Hasil.....	80
6.2. Hasil Praproses Data	80
6.2.1. Hasil <i>Case Folding</i>	80
6.2.2. Hasil Tokenisasi	81
6.3. Hasil Proses <i>Stemming</i>	81
6.4. Hasil Evaluasi Waktu Proses	82
6.5. Hasil Evaluasi Akurasi	84
6.5.1. Evaluasi Kesalahan <i>Stemming ECS</i>	86
6.5.2. Evaluasi Kesalahan <i>Stemming ECS</i> tanpa Kamus dan Porter	89
6.6. Komparasi Hasil Evaluasi Waktu Proses dengan Akurasi.....	91
BAB VII KESIMPULAN	93
7.1. Kesimpulan	93
7.2. Saran.....	93
DAFTAR PUSTAKA	95
LAMPIRAN.....	97

DAFTAR GAMBAR

Gambar 3.1	Ilustrasi Proses <i>Stemming</i> (Purwoko, 2011)	21
Gambar 3.2	Metode <i>Conflation</i> (Sharma, 2012)	22
Gambar 3.3	Tipe dari Algoritma <i>Stemming</i> (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	<i>Pseudocode Binary Search (iterative)</i>	39
Gambar 4.4	Cuplikan Tabel Verifikasi Kata Kunci	41
Gambar 4.5	Diagram Alir Praproses Data	41
Gambar 4.6	Diagram Alir Proses <i>Case Folding</i>	42
Gambar 4.7	Diagram Alir Proses Tokenisasi	42
Gambar 4.8	<i>Pseudocode Enhanced Confix Stripping Stemmer</i> (1)	45
Gambar 4.9	<i>Pseudocode Enhanced Confix Stripping Stemmer</i> (2)	46
Gambar 4.10	Diagram Alir <i>Stemming ECS</i> (1)	46
Gambar 4.11	Diagram Alir <i>Stemming ECS</i> (2)	47
Gambar 4.12	Diagram Alir <i>Stemming ECS</i> (3)	48
Gambar 4.13	<i>Pseudocode ECS Stemmer Tanpa Kamus</i> (1)	50
Gambar 4.14	<i>Pseudocode ECS Stemmer Tanpa Kamus</i> (2)	51
Gambar 4.15	Diagram Alir <i>Stemming ECS Tanpa Kamus</i> (1)	51
Gambar 4.16	Diagram Alir <i>Stemming ECS Tanpa Kamus</i> (2)	52
Gambar 4.17	<i>Pseudocode Porter Stemmer</i>	54
Gambar 4.18	Diagram Alir Porter <i>Stemmer</i> (1)	55
Gambar 4.19	Diagram Alir Porter <i>Stemmer</i> (2)	56
Gambar 4.20	<i>Confusion Matrix</i> (Fawcett, 2006)	58
Gambar 5.1	Kode Program Pengumpulan Data dengan <i>Scraping</i>	62
Gambar 5.2	Cuplikan Data Uji	63
Gambar 5.3	Cuplikan Hasil Validasi Kata Kunci	63
Gambar 5.4	Implementasi <i>Case Folding</i>	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 <i>Singular</i> atau <i>Plural</i>	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 <i>class AbstractDisambiguatePrefixRule</i>	71
Gambar 5.15	Kode Program Pemotongan Sisipan	71
Gambar 5.16	Kode Program <i>loopPengembalian Akhiran</i> (1)	72
Gambar 5.17	Kode Program <i>loopPengembalian Akhiran</i> (2)	73
Gambar 5.18	Cuplikan Kode Program Tahapan <i>Stemming ECS</i> (1)	73
Gambar 5.19	Cuplikan Kode Program Tahapan <i>Stemming ECS</i> (2)	74

Gambar 5.20	Kode Program <i>class DontStemShortWord</i>	74
Gambar 5.21	Cuplikan Kode Program <i>Stemming</i> Porter (1)	75
Gambar 5.22	Cuplikan Kode Program <i>Stemming</i> Porter (2)	76
Gambar 5.23	Cuplikan Kode Program <i>Stemming</i> Porter (3)	77
Gambar 5.24	Format Argumen Pengujian <i>Stemmer</i>	77
Gambar 5.25	Kode Program Implementasi Pengujian <i>ECS Stemmer</i> (1)	77
Gambar 5.26	Kode Program Implementasi Pengujian <i>ECS Stemmer</i> (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 <i>Stemming</i>	79
Gambar 6.1	Cuplikan Data Hasil Proses <i>Case Folding</i>	80
Gambar 6.2	Hasil Proses Tokenisasi	81
Gambar 6.3	Cuplikan Hasil Keluaran dari <i>Stemmer</i>	81
Gambar 6.4	Grafik Rata-rata Waktu Proses <i>Stemmer</i>	83
Gambar 6.5	Grafik Nilai Akurasi <i>Stemmer</i>	85
Gambar 6.6	Grafik Jumlah Kesalahan <i>Stemmer ECS</i>	87
Gambar 6.7	Diagram Evaluasi <i>Overstemming</i> dan <i>Understemming ECS</i>	89
Gambar 6.8	Diagram Evaluasi Tingkat Kesalahan <i>Stemming ECS</i> tanpa Kamus dan Porter	90
Gambar 6.9	Diagram Komparasi Hasil Evaluasi <i>Stemmer</i>	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 <i>Inflectional Particle</i>	32
Tabel 3.5	Aturan Untuk <i>Inflectional Possesive Pronoun</i>	32
Tabel 3.6	Aturan Untuk <i>First Order Derivational Prefix</i>	32
Tabel 3.7	Aturan Untuk <i>Second Order Derivational Prefix</i>	32
Tabel 3.8	Aturan Untuk <i>Derivational Suffix</i>	33
Tabel 3.9	Pola Umum Pada <i>Regex</i>	34
Tabel 3.10	Metode Pencocokan dan Pencarian <i>String</i>	35
Tabel 3.11	Metode Modifikasi <i>String</i>	35
Tabel 4.1	Tabel Pemenggalan Awalan <i>Stemmer ECS</i>	49
Tabel 4.2	Aturan Pemenggalan untuk Porter <i>Stemmer</i>	53
Tabel 4.3	Rancangan Analisis Kesalahan <i>Stemming</i>	60
Tabel 6.1	Hasil Evaluasi Waktu Proses <i>Stemmer</i>	82
Tabel 6.2	Hasil Evaluasi Akurasi <i>Stemmer</i>	84
Tabel 6.3	Evaluasi Kesalahan Hasil <i>Stemming ECS</i>	86
Tabel 6.4	Evaluasi Kesalahan <i>Stemming ECS</i> tanpa Kamus dan Porter (1)	89
Tabel 6.5	Evaluasi Kesalahan <i>Stemming ECS</i> tanpa Kamus dan Porter (2)	90