BAB 4
IMPLEMENTASI DAN PENGUJIAN

4.1 Implementasi Sistem

Tahap implementasi sistem menjelaskan pengimplementasian dari simulator yang akan dibuat. Tahapan yang dilakukan untuk menerjemahkan perancangan berdasarkan hasil analisis dalam bahasa yang dimengerti komputer serta penerapan perangkat lunak pada keadaan yang sebenarnya. Implementasi yang akan dijelaskan berupa implementasi perangkat keras, implementasi perangkat lunak, implementasi basis data dan implementasi antarmuka.

4.1.1 Implementasi Perangkat Keras

Implementasi perangkat keras menjelaskan tentang perangkat keras yang digunakan untuk membangun simulator yang telah dibuat. Pada tabel 4.1 dibawah ini dapat dilihat perangkat keras yang digunakan dalam simulator ini.

Tabel 4.1 Implementasi Perangkat Keras

No	Perangkat Keras	Spesifikasi
1	Prosessor	Prosessor Core i3 @1.8 GHz
2	RAM	2GB
3	Monitor	LCD 14”
4	VGA	512 MB
5	*Harddisk*	160GB
6	Peranti pendukung	*Keyboard, mouse* dan modem indihome

4.1.2 Implementasi Perangkat Lunak

Implementasi Perangkat lunak menjelaskan tentang perangkat lunak yang akan digunakan untuk membangun simulator yang telah dibuat. Adapun perangkat lunak yang digunakan dalam implementasi simulator dapat dilihat pada tabel 4.2 dibawah ini.

Tabel 4.2 Implementasi Perangkat Lunak

No	Perangkat Lunak
1	Sistem Operasi Windows XP/7

2	Microsoft Visual Studio 2010
3	Wampserver
4	Notepad ++
5	MySQL

4.1.3 Implementasi Basis Data

Pembuatan basis data dalam simulator ini dilakukan dengan menggunakan MySQL. Implementasi basis data dalam bahasa SQL dapat dilihat pada table 4.3 dibawah ini.

Tabel 4.3 Implementasi Basis Data

No	Proses	SQL
1	Proses Pembuatan Database	CREATE DATABASE db_skripsi;
2	Tabel “tb_katadasar”	CREATE TABLE ‘tb_katadasar’ (id_katadasar int(10) NOT NULL AUTO_INCREMENT, katadasar varchar(70) NOT NULL, tipe_katadasar varchar(25) NOT NULL, PRIMARY KEY (‘id_katadasar’)) ENGINE=InnoDB;

4.1.4 Implementasi Antarmuka

Implementasi antarmuka dilakukan untuk mengimplementasikan setiap tampilan yang terdapat pada tahap perancangan yang telah dilakukan. Adapun implementasi antarmuka dari simulator yang telah dibuat dapat dilihat pada tabel 4.4 dibawah ini.

Tabel 4.4 Implementasi Antarmuka

Nama Antarmuka	Proses	Nama File	Keterangan
Halaman utama	Tampilan awal	FormMenu.cs	Tampilan awal saat memasuki simulator
Form *Stemming*	Proses metode *stemming*	FormStemmer.cs	Tampilan pada saat dilakukannya proses membandungkan dua

			metode *stemming* yaitu algoritma Idris dan Algoritma ECS *Stemmer*

4.2 Pengujian Metode

Tahapan pengujian yang akan dilakukan terdiri dari pengujian dari masing – masing algoritma *stemming* yaitu Algoritma Idris dan Algoritma ECS *Stemmer*. Pengujian ini dilakukan dengan tujuan untuk mengetahui performansi dari masing – masing algoritma Idris dan algoritma ECS *Stemmer* yang baik digunakan untuk menstemming kata baku dalam bahasa Indonesia yang akan dilakukan pada simulator yang telah dibuat, pengujian meliputi akurasi dan waktu proses yang akan dilakukan untuk menghasilkan *output*. Skenario pengujian sangat diperlukan agar proses pengujian yang dilakukan dapat mencapai tujuan yang diinginkan. Adapun skenario pengujian yang dilakukan yaitu melakukan proses pencarian kata dasar pada *database* dan menganalisis hasil *stemming*[4]. Berikut deskripsi mengenai skenario pengujian yang dilakukan adalah sebagai berikut.

a. Skenario Pengujian 1

Menguji *stemmer strength* dari Algoritma Idris dan Algoritma ECS *Stemmer*. Pengujian dilakukan dengan menganalisis nilai yang dihasilkan berupa nilai jumlah kata sebelum dan sesudah *stemming* dengan nilai icf dan persentase. Skenario yang dilakukan yaitu dengan menggunakan 10 dokumen [5] dan menganalisis nilai parameter yang dihasilkan. Berikut adalah skenario pengujian pertama yang dilakukan.

1. Menggunakan kata-kata sesuai dengan dokumen yang telah ditentukan. Pada pengujian skenario 1 ini, satu dokumen diuji dua kali karena pada simulator ini apabila melakukan *stemming* pengguna memilih salah satu *stemming* yang akan diuji terlebih dahulu. Oleh karena itu pada pengujian dokumen pertama dipilih algoritma Idris terlebih dahulu kemudian baru dipilih algoritma ECS *Stemmer*. Pada pengujian kedua dengan dokumen

yang sama dipilih algoritma ECS *Stemmer* terlebih dahulu kemudian memilih algoritma Idris hal ini kebalikan dari pengujian pertama.

2. Menghasilkan *output* berupa kata dasar setelah proses *stemming*.
3. Menghitung nilai parameter icf dan persentase dari masing-masing dokumen.

b. Skenario Pengujian 2

Menguji keakuratan Algoritma Idris dan Algoritma ECS *Stemmer* terhadap stem yang dihasilkan. Skenario yang dilakukan, yakni menggunakan 10 dokumen yang telah ditentukan sebelumnya dan menganalisis isi dokumen uji berdasarkan jumlah kata yang tidak berhasil distemming. Berikut adalah skenario pengujian kedua yang dilakukan.

1. Menggunakan kata-kata sesuai dokumen yang telah ditentukan.
2. Menghasilkan *output* berupa kata-kata setelah dilakukan proses *stemming*.
3. Menganalisis dan menghitung jumlah kata yang tidak berhasil distemming dengan benar.

4.2.1 Skenario Pengujian 1

1. Pengujian dokumen ke – 1

Berikut ini adalah hasil analisis yang telah dilakukan.

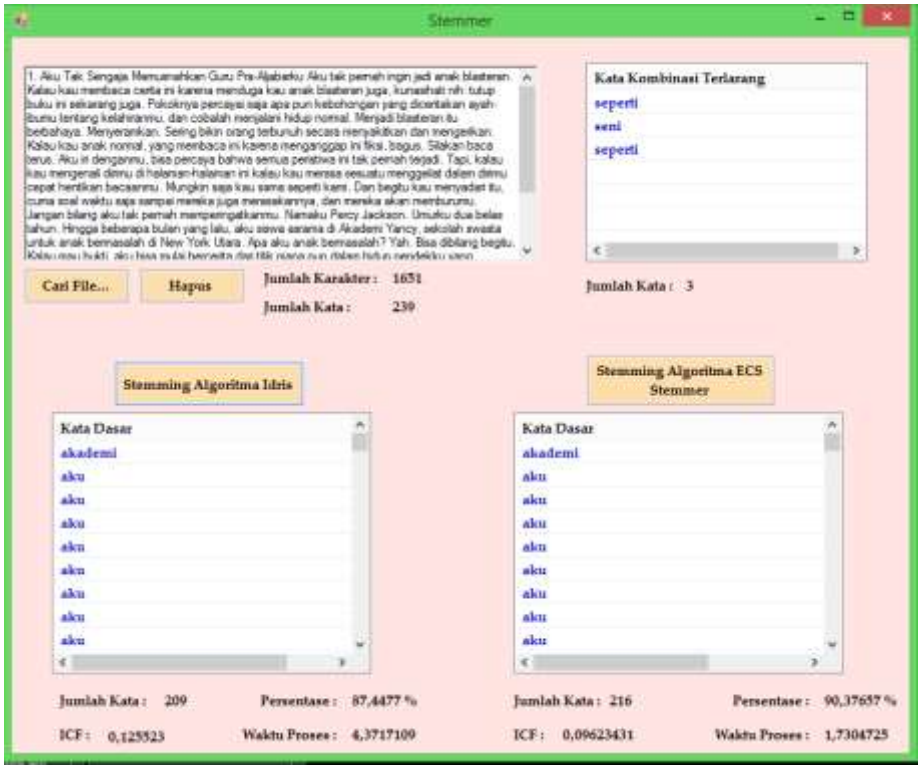
Nama dokumen : Percy Jackson dan Pencuri Petir

Jumlah karakter : 1651 karakter

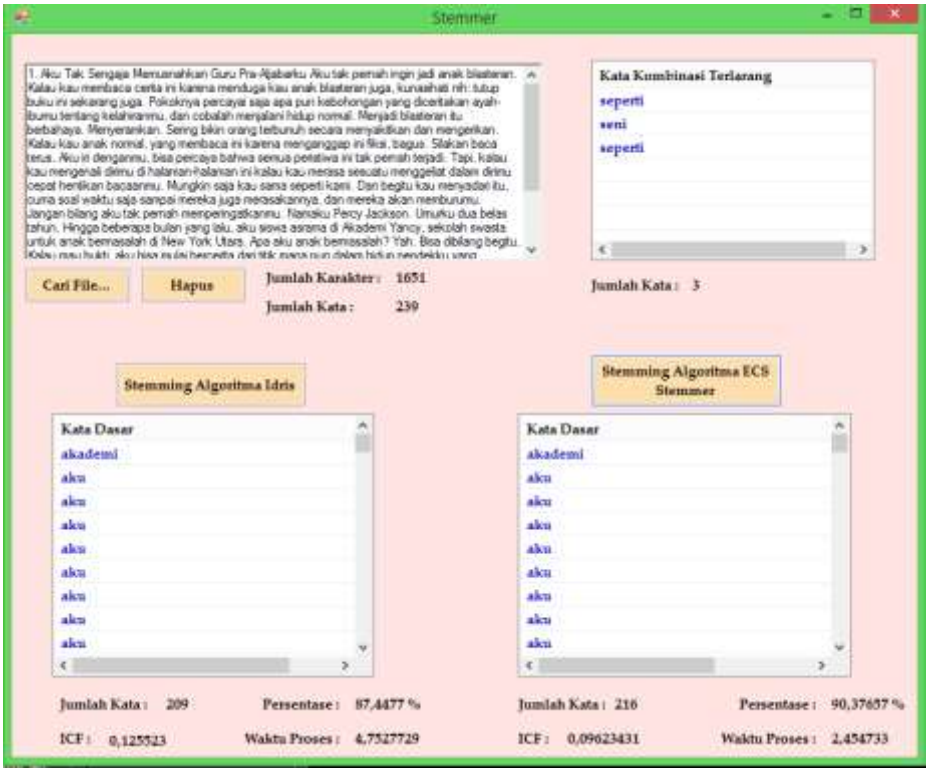
Jumlah kata : 239 kata

Ukuran dokumen : 2 kb

Hasil Pengujian dari dokumen ke 1 dapat dilihat pada gambar 4.1 dan gambar 4.2



Gambar 4.1 Hasil pengujian dokumen 1.1



Gambar 4.2 Hasil pengujian dokumen 1.2

Pada pengujian dokumen ke 1.1 dengan ukuran dokumen sebesar 2 kb dan jumlah kata terdiri dari 239 kata, pada gambar 4.1 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 209 kata dengan ICF sebesar 0,125523, persentase kekuatan stemmernya 87% dan waktu pemrosesan 4,37 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 216 kata dengan ICF sebesar 0,09623431, persentase kekuatan stemmernya sebesar 90% dan waktu pemrosesannya 1,73 detik.

Pada pengujian dokumen 1.2 dengan ukuran sebesar 2 kb dan jumlah kata yang terdiri dari 239 kata, pada gambar 4.2 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 209 kata dengan ICF sebesar 0,125523, persentase kekuatan stemmernya 87% dan waktu pemrosesan 4,75 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 216 kata dengan ICF sebesar 0,09623431, persentase kekuatan stemmernya sebesar 90% dan waktu pemrosesannya 2,45 detik.

2. Pengujian dokumen ke – 2

Berikut ini adalah hasil analisis yang telah dilakukan

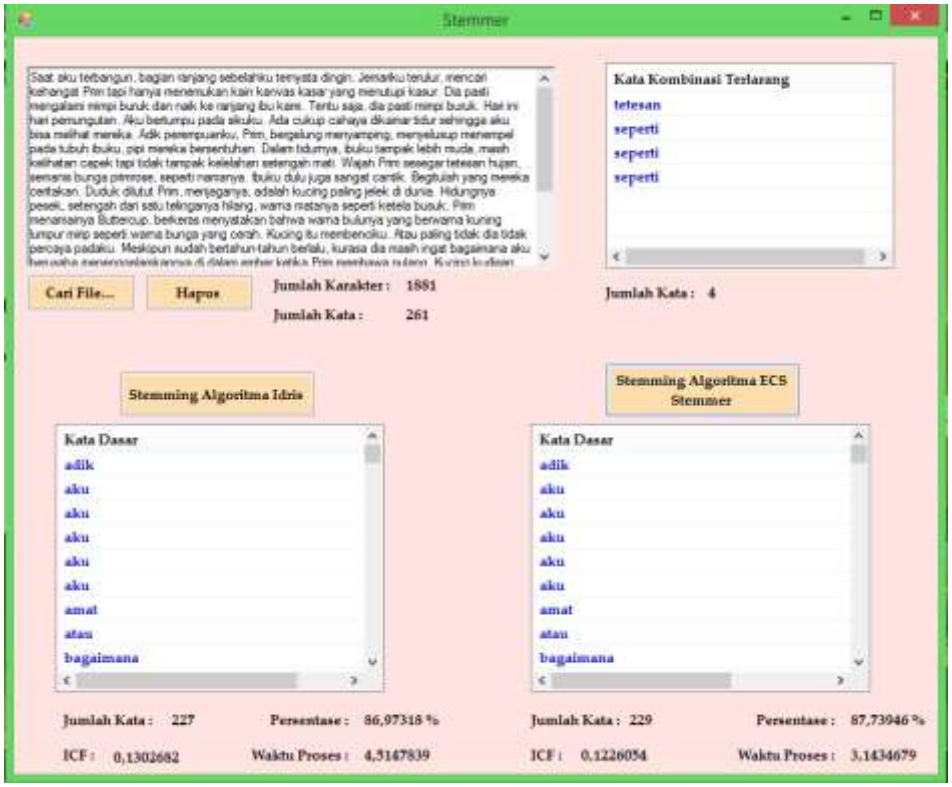
Nama dokumen : The Hunger Games

Jumlah karakter : 1881 karakter

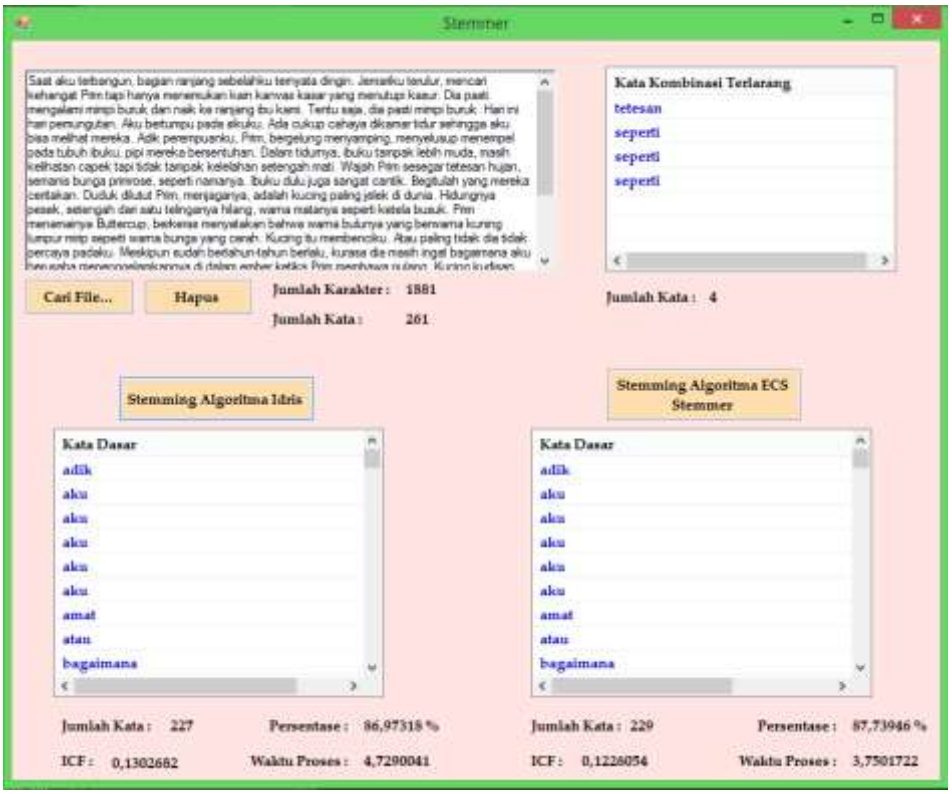
Jumlah kata : 261 kata

Ukuran dokumen : 2 kb

Hasil Pengujian dari dokumen ke 2 dapat dilihat pada gambar 4.3 dan gambar 4.4



Gambar 4.3 Hasil pengujian dokumen 2.1



Gambar 4.4 Hasil pengujian dokumen 2.2

Pada pengujian dokumen ke 2.1 dengan ukuran dokumen sebesar 2 kb dan jumlah kata terdiri dari 261 kata, pada gambar 4.3 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 227 kata dengan ICF sebesar 0,1302682, persentase kekuatan stemmernya 86% dan waktu pemrosesan 4,51 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 229 kata dengan ICF sebesar 0,1226054, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 3,14 detik.

Pada pengujian dokumen 2.2 dengan ukuran sebesar 2 kb dan jumlah kata yang terdiri dari 239 kata, pada gambar 4.4 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 227 kata dengan ICF sebesar 0,1302682, persentase kekuatan stemmernya 86% dan waktu pemrosesan 4,72 detik sedangkan dengan Algoritma ECS *Stemmer*, kata yang berhasil distemming sebanyak 229 kata dengan ICF sebesar 0,1226054, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 3,75 detik.

3. Pengujian dokumen ke – 3

Berikut ini adalah hasil analisis yang telah dilakukan.

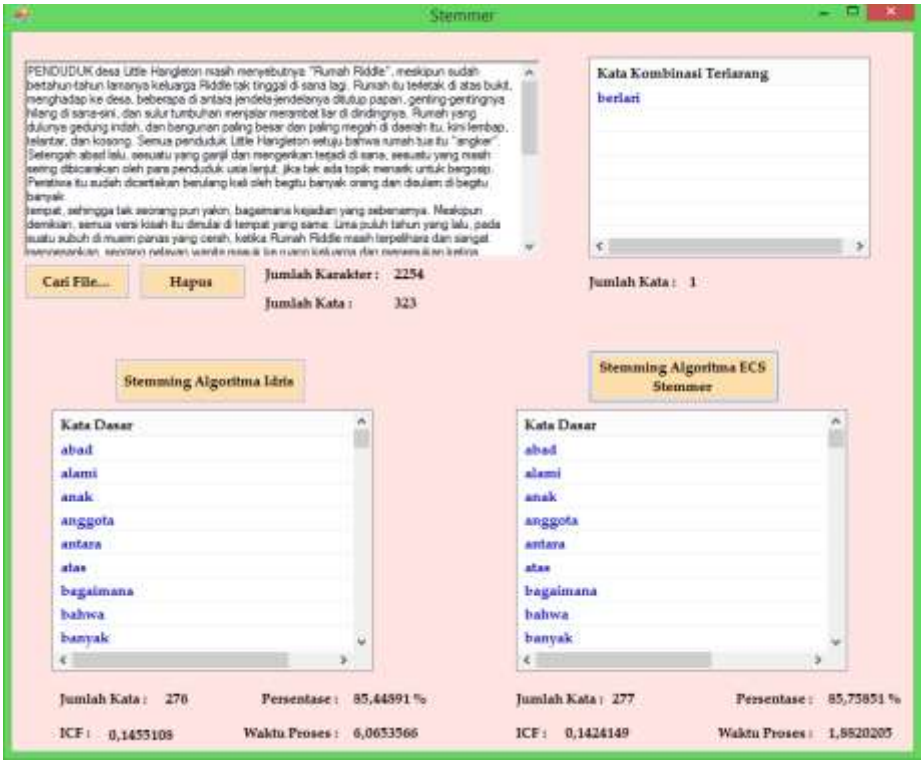
Nama dokumen : Harry Potter dan Piala Api

Jumlah karakter : 2254 karakter

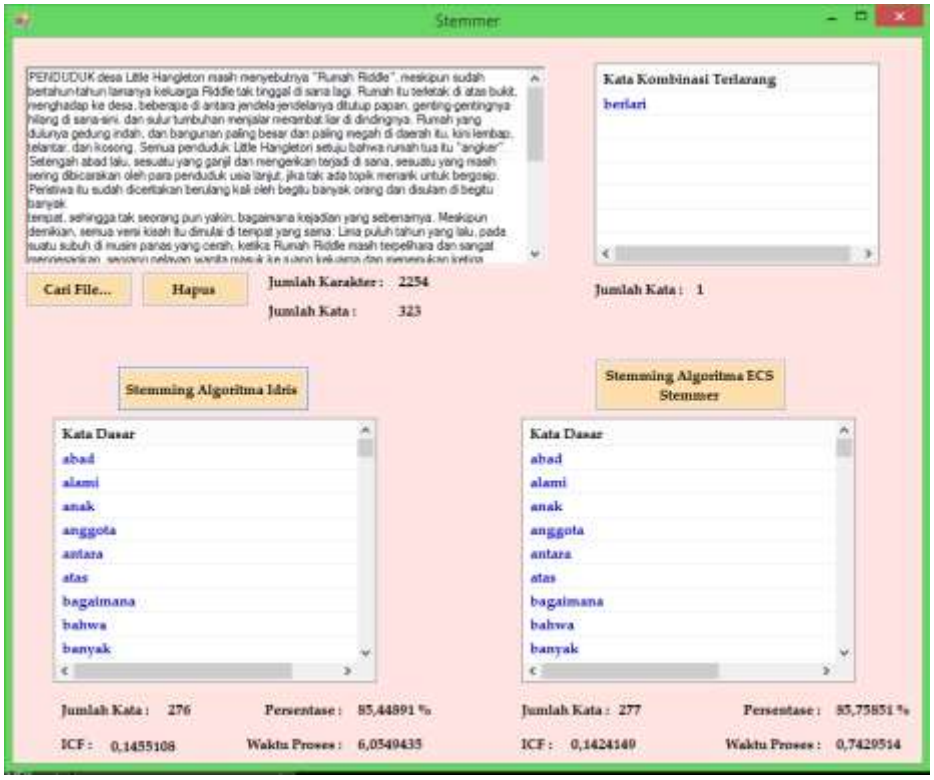
Jumlah kata : 323 kata

Ukuran dokumen : 3 kb

Hasil Pengujian dari dokumen ke 3 dapat dilihat pada gambar 4.5 dan gambar 4.6.



Gambar 4.5 Hasil pengujian dokumen 3.1



Gambar 4.6 Hasil pengujian dokumen 3.2

Pada pengujian dokumen ke 2.1 dengan ukuran dokumen sebesar 2 kb dan jumlah kata terdiri dari 261 kata, pada gambar 4.3 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 227 kata dengan ICF sebesar 0,1302682, persentase kekuatan stemmernya 86% dan waktu pemrosesan 4,51 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 229 kata dengan ICF sebesar 0,1226054, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 3,14 detik.

Pada pengujian dokumen 2.2 dengan ukuran sebesar 2 kb dan jumlah kata yang terdiri dari 239 kata, pada gambar 4.4 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 227 kata dengan ICF sebesar 0,1302682, persentase kekuatan stemmernya 86% dan waktu pemrosesan 4,72 detik sedangkan dengan Algoritma ECS *Stemmer*, kata yang berhasil distemming sebanyak 229 kata dengan ICF sebesar 0,1226054, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 3,75 detik.

4. Pengujian dokumen ke – 4

Berikut ini adalah hasil analisis yang telah dilakukan.

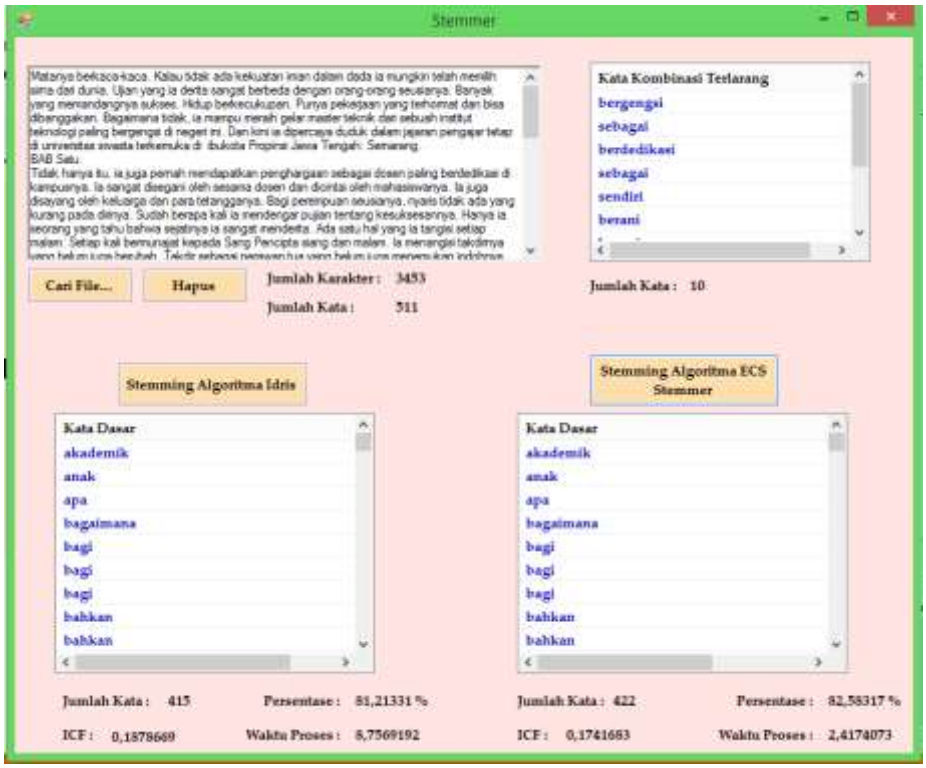
Nama dokumen : Dalam Mihrab Cinta

Jumlah karakter : 3453 karakter

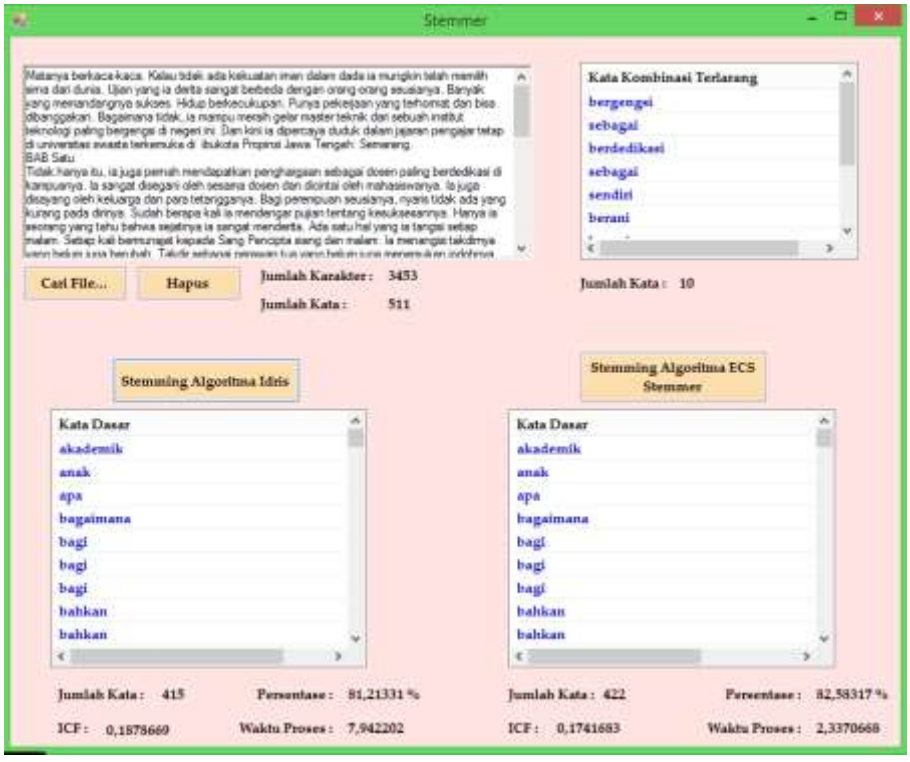
Jumlah kata : 511 kata

Ukuran dokumen : 4 kb

Hasil Pengujian dari dokumen ke 4 dapat dilihat pada gambar 4.7 dan gambar 4.8



Gambar 4.7 Hasil pengujian dokumen 4.1



Gambar 4.8 Hasil pengujian dokumen 4.2

Pada pengujian dokumen ke 4.1 dengan ukuran dokumen sebesar 4 kb dan jumlah kata terdiri dari 511 kata, pada gambar 4.7 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 415 kata dengan ICF sebesar 0,1878669, persentase kekuatan stemmernya 81% dan waktu pemrosesan 8,75 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 422 kata dengan ICF sebesar 0,1741683, persentase kekuatan stemmernya sebesar 82% dan waktu pemrosesannya 2,41 detik.

Pada pengujian dokumen 4.2 dengan ukuran sebesar 4 kb dan jumlah kata yang terdiri dari 511 kata, pada gambar 4.8 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 415 kata dengan ICF sebesar 0,1878669, persentase kekuatan stemmernya 81% dan waktu pemrosesan 7,94 detik sedangkan dengan Algoritma ECS *Stemmer*, kata yang berhasil distemming sebanyak 422 kata dengan ICF sebesar 0,1741683, persentase kekuatan stemmernya sebesar 82% dan waktu pemrosesannya 2,33 detik.

5. Pengujian dokumen ke – 5

Berikut ini adalah hasil analisis yang telah dilakukan

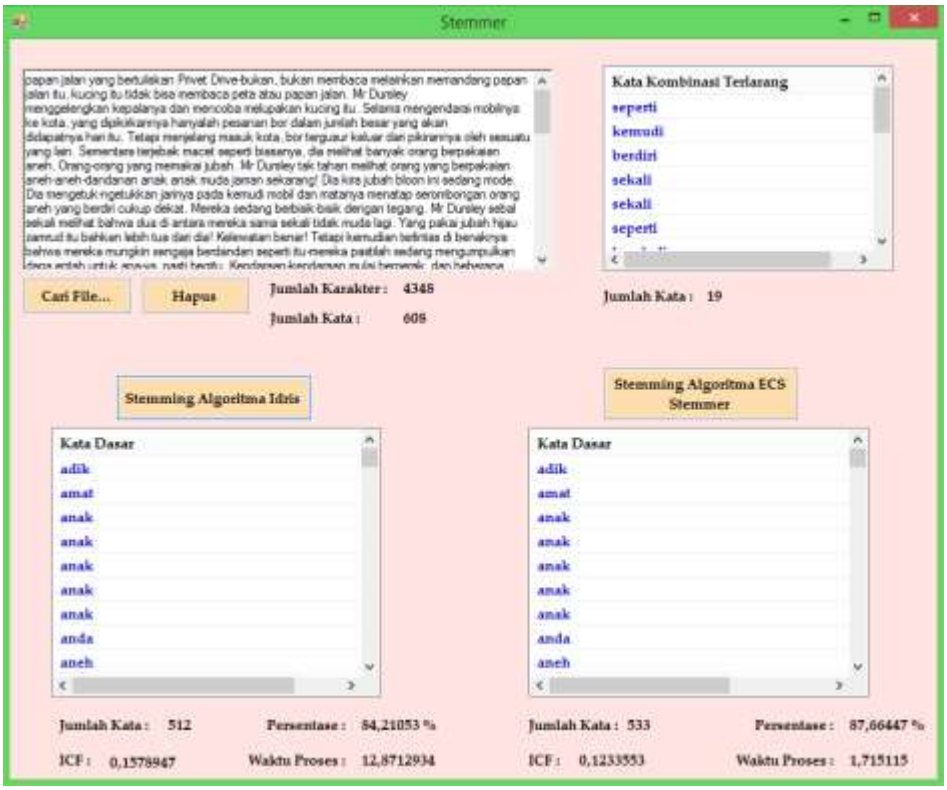
Nama dokumen : Harry Potter dan Batu Bertuah

Jumlah karakter : 4348 karakter

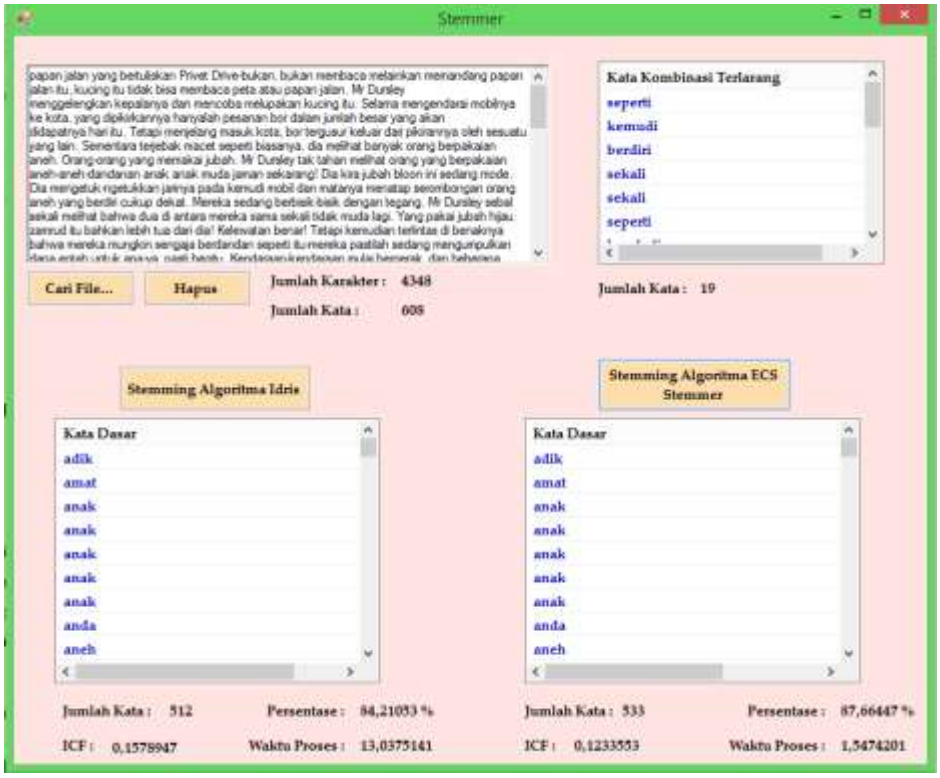
Jumlah kata : 608 kata

Ukuran dokumen : 5 kb

Hasil Pengujian dari dokumen ke 5 dapat dilihat pada gambar 4.9 dan gambar 4.10



Gambar 4.9 Hasil pengujian dokumen 5.1



Gambar 4.10 Hasil pengujian dokumen 5.2

Pada pengujian dokumen ke 5.1 dengan ukuran dokumen sebesar 5 kb dan jumlah kata terdiri dari 608 kata, pada gambar 4.9 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 512 kata dengan ICF sebesar 0,1578947, persentase kekuatan stemmernya 84% dan waktu pemrosesan 12,87 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 533 kata dengan ICF sebesar 0,1233553, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 1,71 detik.

Pada pengujian dokumen 5.2 dengan ukuran sebesar 5 kb dan jumlah kata yang terdiri dari 511 kata, pada gambar 4.10 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 512 kata dengan ICF sebesar 0,1578947, persentase kekuatan stemmernya 84% dan waktu pemrosesan 13,03 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 533 kata dengan ICF sebesar 0,1233553, persentase kekuatan stemmernya sebesar 87% dan waktu pemrosesannya 1,54 detik.

6. Pengujian dokumen ke – 6

Berikut ini adalah hasil analisis yang telah dilakukan

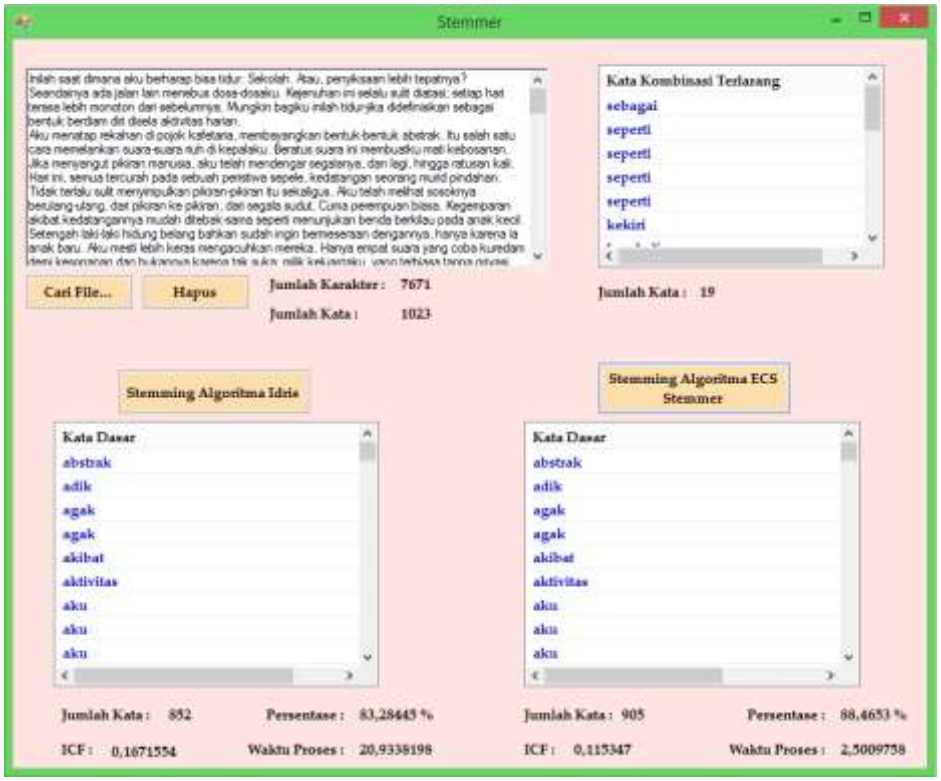
Nama dokumen : Twilight – Midnight Sun

Jumlah karakter : 7671 karakter

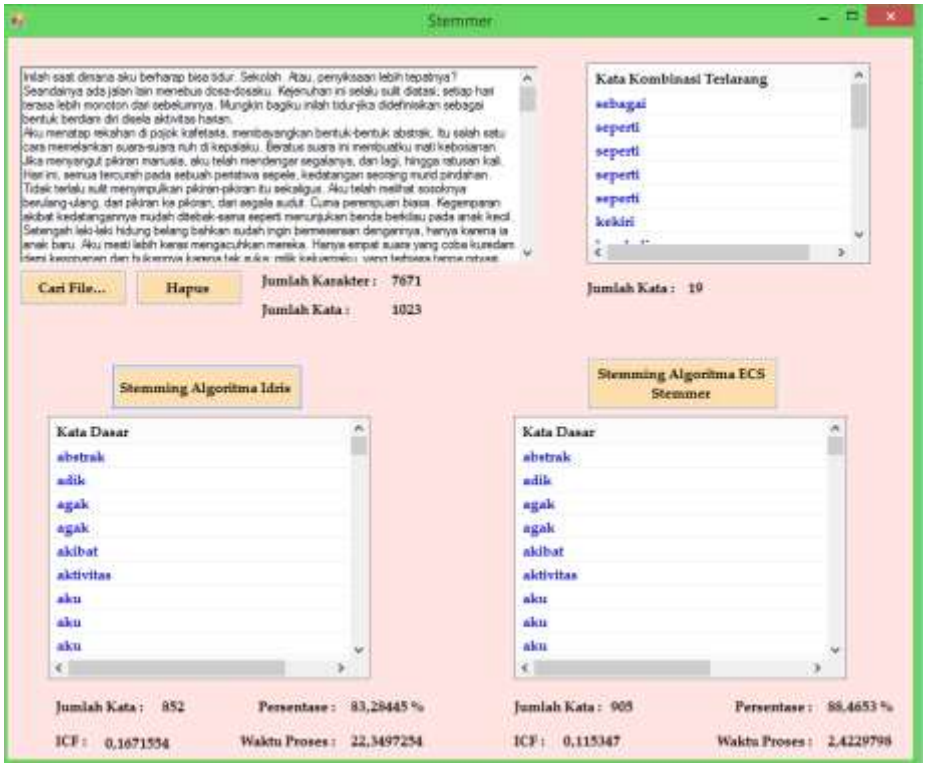
Jumlah kata : 1023 kata

Ukuran dokumen : 8 kb

Hasil Pengujian dari dokumen ke 6 dapat dilihat pada gambar 4.11 dan gambar 4.12



Gambar 4.11 Hasil pengujian dokumen 6.1



Gambar 4.12 Hasil pengujian dokumen 6.2

Pada pengujian dokumen ke 6.1 dengan ukuran dokumen sebesar 8 kb dan jumlah kata terdiri dari 1023 kata, pada gambar 4.11 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 852 kata dengan ICF sebesar 0,1671554, persentase kekuatan stemmernya 83% dan waktu pemrosesan 20,93 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 905 kata dengan ICF sebesar 0,115347, persentase kekuatan stemmernya sebesar 88% dan waktu pemrosesannya 2,50 detik.

Pada pengujian dokumen 6.2 dengan ukuran sebesar 8 kb dan jumlah kata yang terdiri dari 1023 kata, pada gambar 4.12 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 852 kata dengan ICF sebesar 0,1671554, persentase kekuatan stemmernya 83% dan waktu pemrosesan 22,34 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 905 kata dengan ICF sebesar 0,115347, persentase kekuatan stemmernya sebesar 88% dan waktu pemrosesannya 2,42 detik.

7. Pengujian dokumen ke – 7

Berikut ini adalah hasil analisis yang telah dilakukan

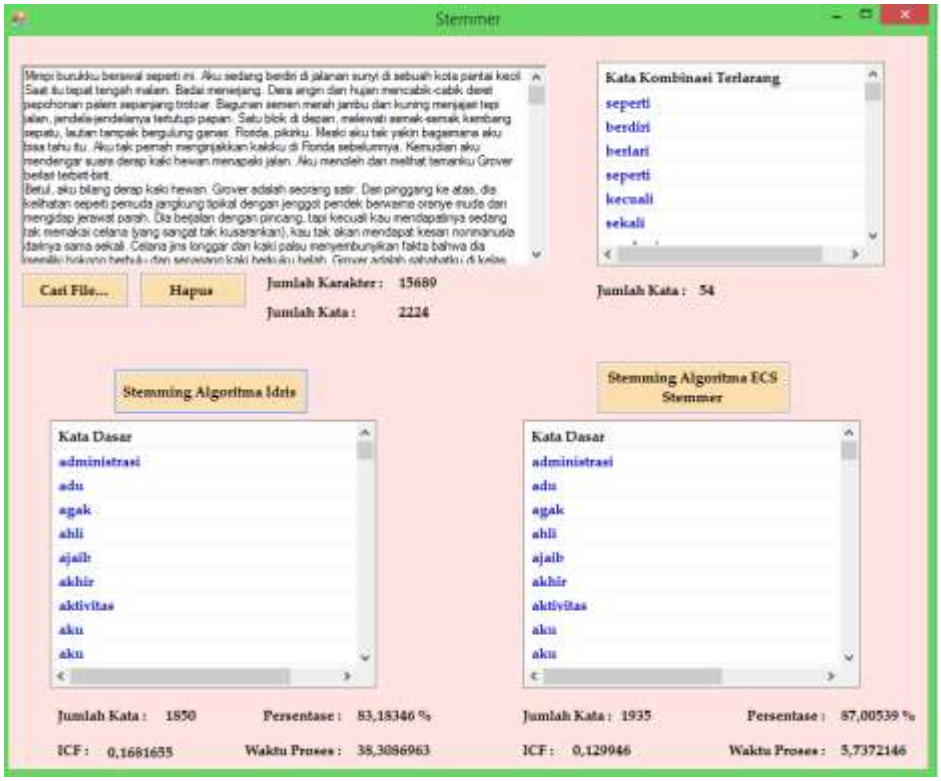
Nama dokumen : Percy Jackson dan Lautan Monster

Jumlah karakter : 15689 karakter

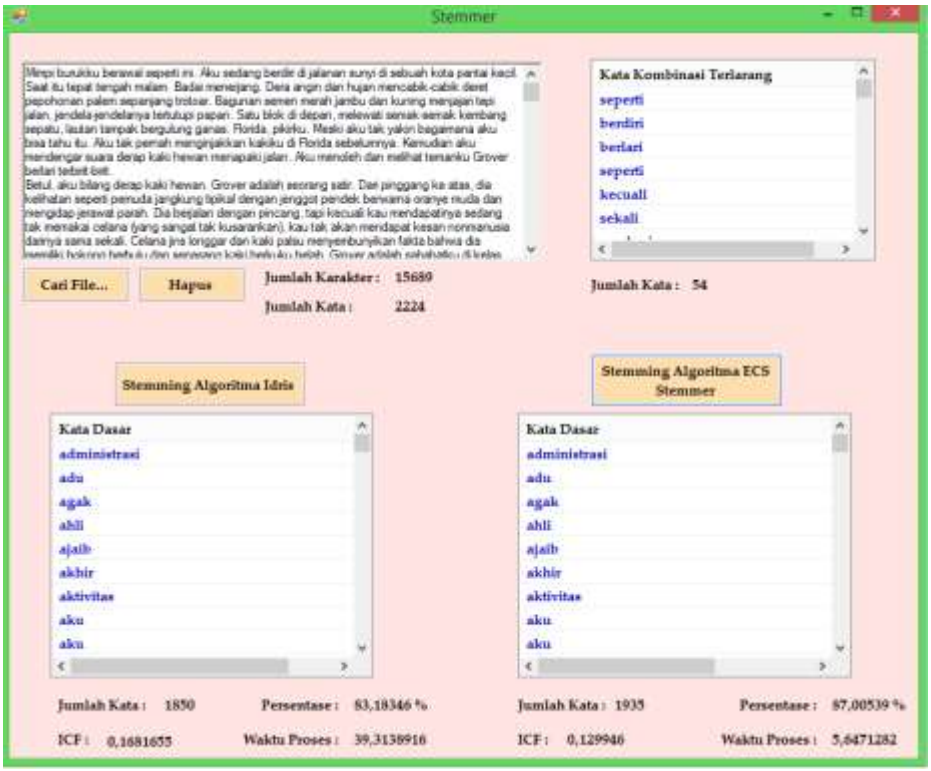
Jumlah kata : 2224 Kata

Ukuran dokumen : 16 kb

Hasil Pengujian dari dokumen ke 7 dapat dilihat pada gambar 4.13 dan gambar 3.14



Gambar 4.13 Hasil pengujian dokumen 7.1



Gambar 4.14 Hasil pengujian dokumen 7.2

Pada pengujian dokumen ke 7.1 dengan ukuran dokumen sebesar 16 kb dan jumlah kata terdiri dari 2224 kata, pada gambar 4.13 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 1850 kata dengan ICF sebesar 0,1681655, persentase kekuatan stemmernya 83% dan waktu pemrosesan 38,30 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 1935 kata dengan ICF sebesar 0,129946, persentase kekuatan stemmernya sebesar 88% dan waktu pemrosesannya 5,73 detik.

Pada pengujian dokumen 7.2 dengan ukuran sebesar 16 kb dan jumlah kata yang terdiri dari 2224 kata, pada gambar 4.14 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 1850 kata dengan ICF sebesar 0,1681655, persentase kekuatan stemmernya 83% dan waktu pemrosesan 39,31 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 1935 kata dengan ICF sebesar 0,129946, persentase kekuatan stemmernya sebesar 88% dan waktu pemrosesannya 5,64 detik.

8. Pengujian dokumen ke – 8

Berikut ini adalah hasil analisis yang telah dilakukan

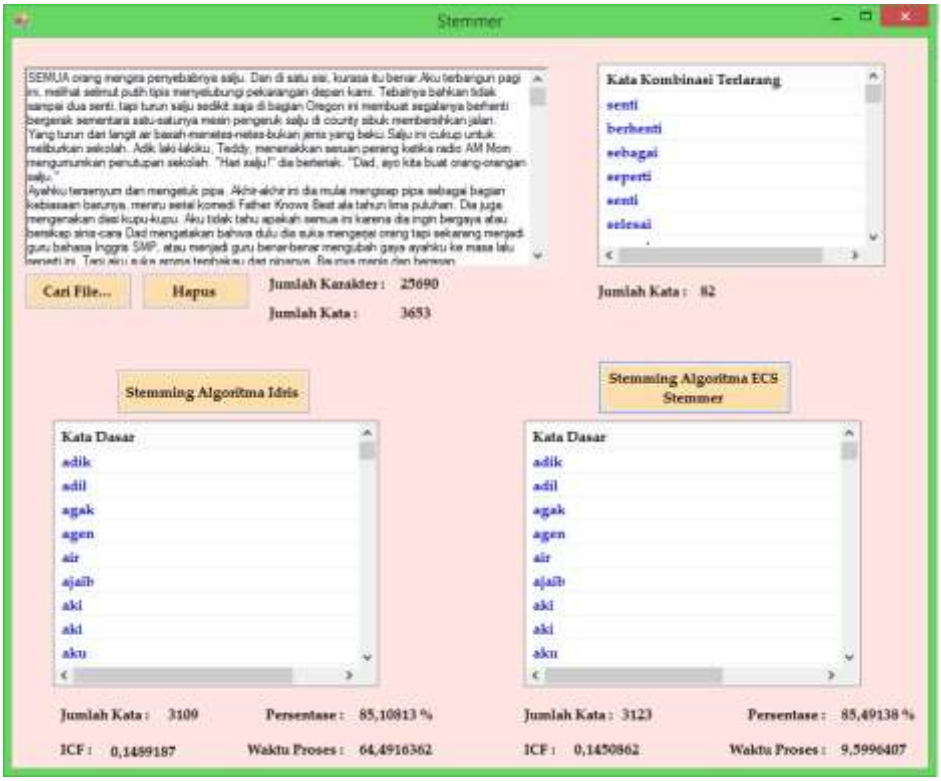
Nama dokumen : If I Stay

Jumlah karakter : 25690 karakter

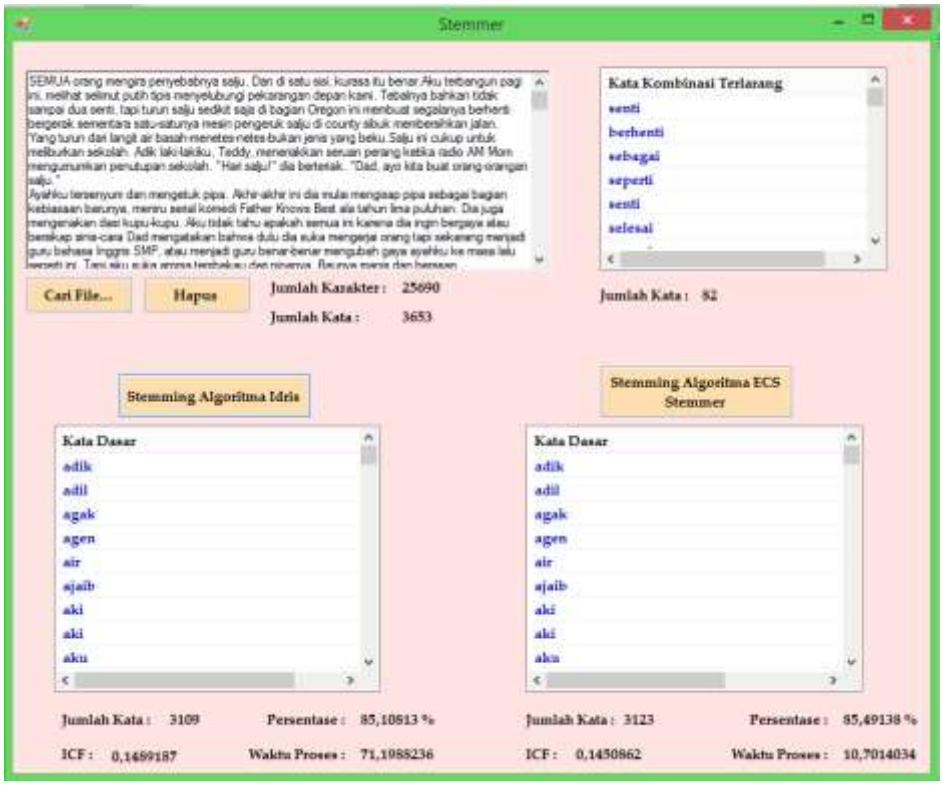
Jumlah kata : 3564 kata

Ukuran dokumen : 26 kb

Hasil Pengujian dari dokumen ke 8 dapat dilihat pada gambar 4.15 dan gambar 4.16



Gambar 4.15 Hasil pengujian dokumen 8.1



Gambar 4.16 Hasil pengujian dokumen 8.2

Pada pengujian dokumen ke 8.1 dengan ukuran dokumen sebesar 26 kb dan jumlah kata terdiri dari 3653 kata, pada gambar 4.15 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 3109 kata dengan ICF sebesar 0,1489187, persentase kekuatan stemmernya 85% dan waktu pemrosesan 64,49 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 3123 kata dengan ICF sebesar 0,11450862, persentase kekuatan stemmernya sebesar 85% dan waktu pemrosesannya 9,59 detik.

Pada pengujian dokumen 8.2 dengan ukuran sebesar 26 kb dan jumlah kata yang terdiri dari 3653 kata, pada gambar 4.16 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 3109 kata dengan ICF sebesar 0,1489187, persentase kekuatan stemmernya 85% dan waktu pemrosesan 71,19 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 3123 kata dengan ICF sebesar 0,11450862, persentase kekuatan stemmernya sebesar 85% dan waktu pemrosesannya 10,70 detik.

9. Pengujian dokumen ke – 9

Berikut ini adalah hasil analisis yang telah dilakukan

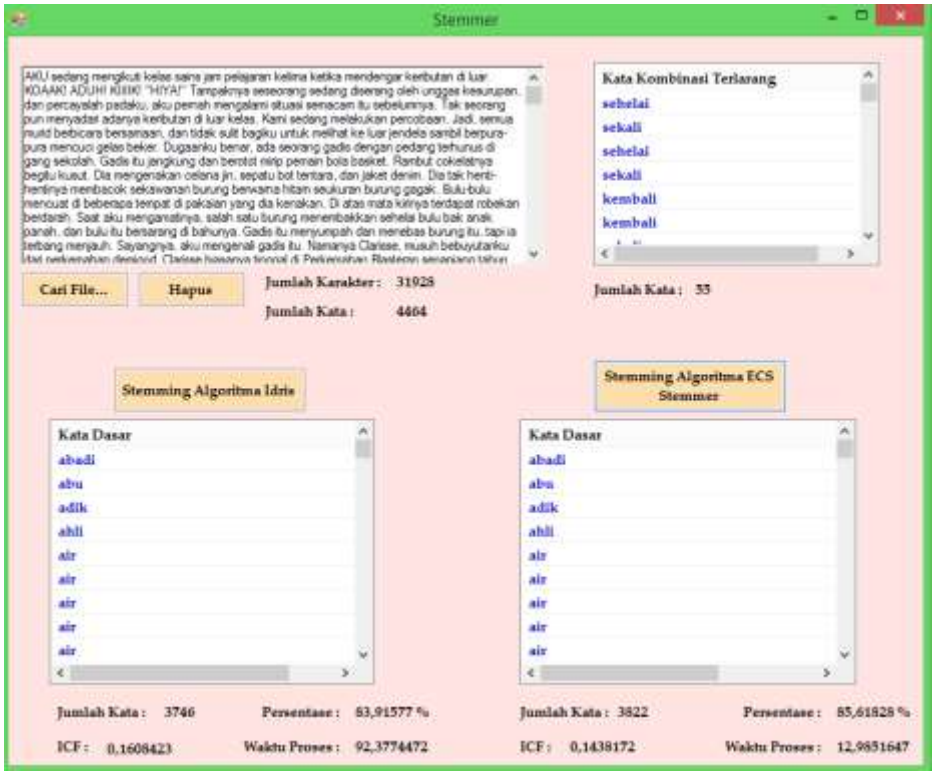
Nama dokumen : The Demigod Diaries

Jumlah karakter : 31928 karakter

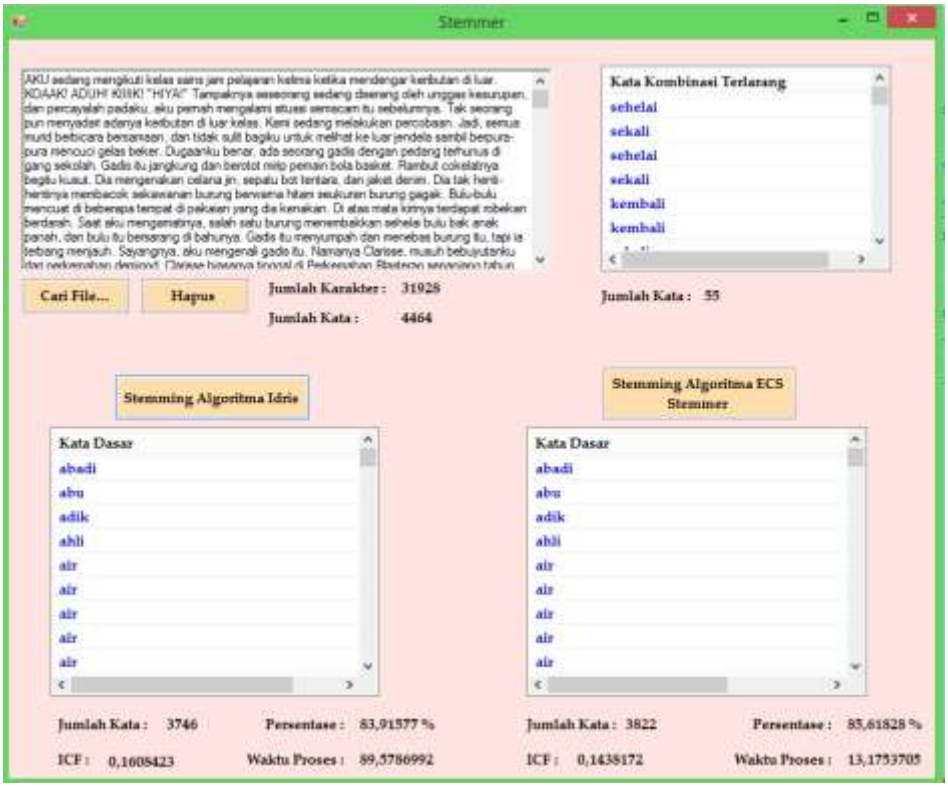
Jumlah kata : 4464 kata

Ukuran dokumen : 32 kb

Hasil Pengujian dari dokumen ke 9 dapat dilihat pada gambar 4.17 dan gambar 4.18



Gambar 4.17 Hasil pengujian dokumen 9.1



Gambar 4.18 Hasil pengujian dokumen 9.2

Pada pengujian dokumen ke 9.1 dengan ukuran dokumen sebesar 32 kb dan jumlah kata terdiri dari 4464 kata, pada gambar 4.17 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 3746 kata dengan ICF sebesar 0,1608423, persentase kekuatan stemmernya 83% dan waktu pemrosesan 92,37 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 3822 kata dengan ICF sebesar 0,1438172, persentase kekuatan stemmernya sebesar 85% dan waktu pemrosesannya 12,98 detik.

Pada pengujian dokumen 9.2 dengan ukuran sebesar 32 kb dan jumlah kata yang terdiri dari 4464 kata, pada gambar 4.18 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 3746 kata dengan ICF sebesar 0,1608423, persentase kekuatan stemmernya 83% dan waktu pemrosesan 89,55 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 3822 kata dengan ICF sebesar 0,1438172, persentase kekuatan stemmernya sebesar 85% dan waktu pemrosesannya 13,17 detik.

10. Pengujian dokumen ke – 10

Berikut ini adalah hasil analisis yang telah dilakukan

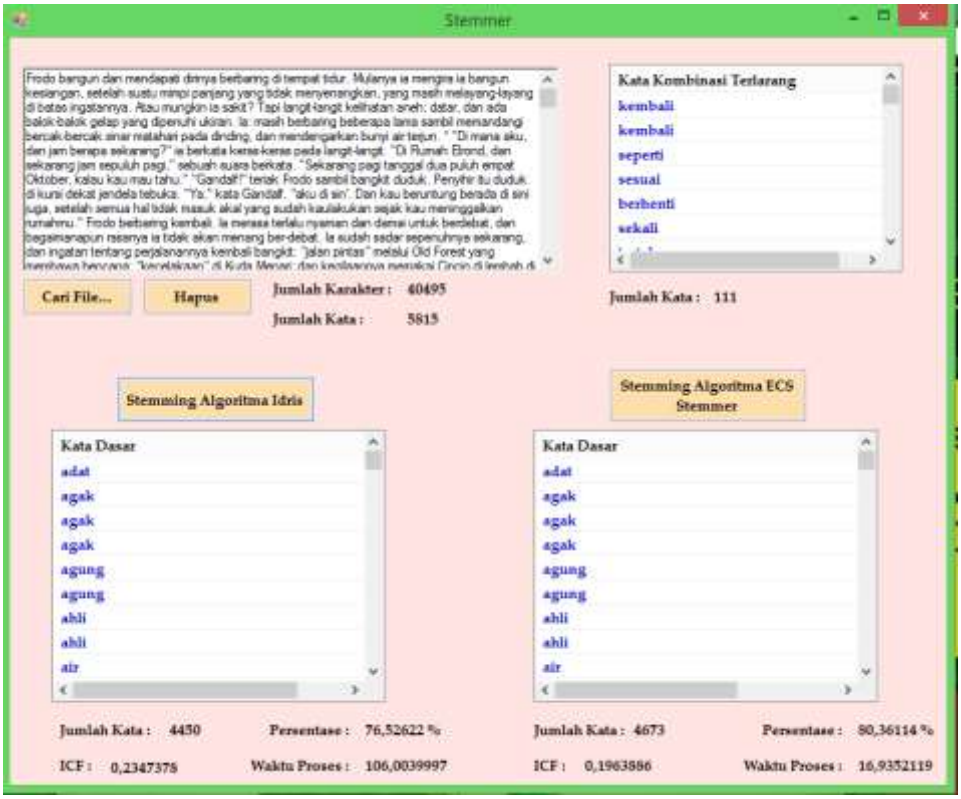
Nama dokumen : The Lord of The Rings – Pembawa Sembilan Cincin

Jumlah karakter : 40495 karakter

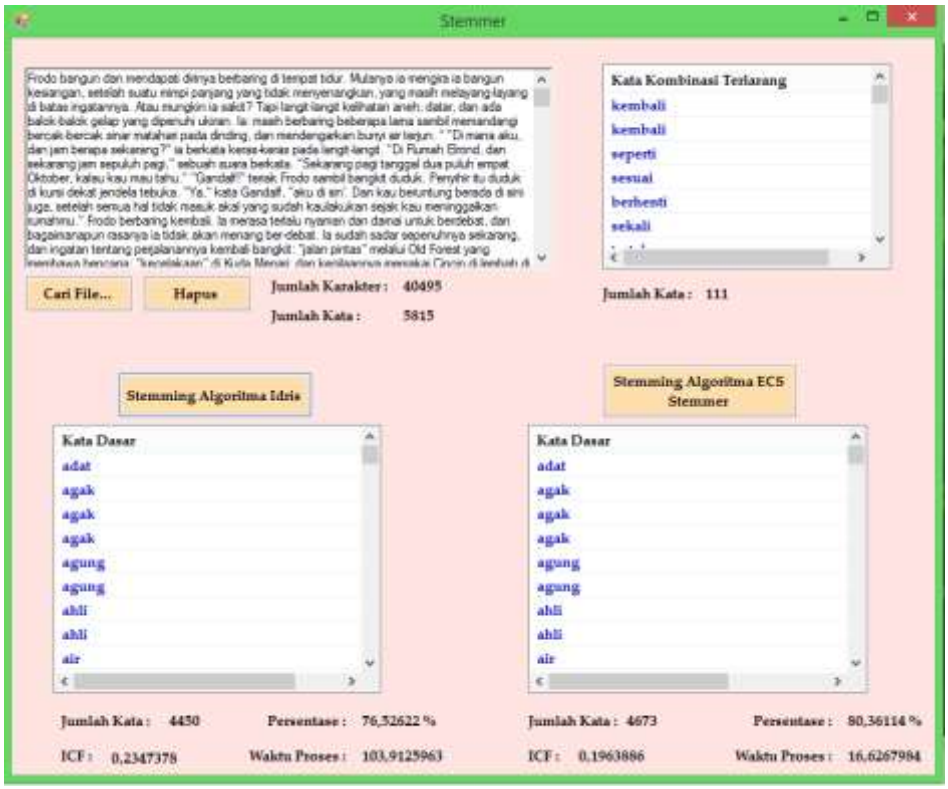
Jumlah kata : 5815 Kata

Ukuran dokumen : 41 kb

Hasil Pengujian dari dokumen ke 10 dapat dilihat pada gambar 4.19 dan gambar 4.20.



Gambar 4.19 Hasil pengujian dokumen 10.1



Gambar 4.20 Hasil pengujian dokumen 10.2

Pada pengujian dokumen ke 10.1 dengan ukuran dokumen sebesar 41 kb dan jumlah kata terdiri dari 5815 kata, pada gambar 4.19 dilakukan pengujian dengan memilih *stemming* dengan algoritma Idris terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma ECS *Stemmer* dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 4450 kata dengan ICF sebesar 0,2347378, persentase kekuatan stemmernya 76% dan waktu pemrosesan 106 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 4673 kata dengan ICF sebesar 0,1963886, persentase kekuatan stemmernya sebesar 80% dan waktu pemrosesannya 16,93 detik.

Pada pengujian dokumen 10.2 dengan ukuran sebesar 41 kb dan jumlah kata yang terdiri dari 5815 kata, pada gambar 4.20 dilakukan pengujian dengan memilih *stemming* dengan algortima ECS *Stemmer* terlebih dahulu kemudian setelah hasil dari *stemming* sudah muncul maka dilakukan *stemming* kedua dengan memilih algoritma Idris dapat dilihat bahwa jumlah kata yang berhasil distemming dengan Algoritma Idris adalah sebanyak 4450 kata dengan ICF sebesar 0,2347378, persentase kekuatan stemmernya 76% dan waktu pemrosesan 103,91 detik sedangkan dengan Algoritma ECS *Stemmer,* kata yang berhasil distemming sebanyak 4673 kata dengan ICF sebesar 0,1963886, persentase kekuatan stemmernya sebesar 80% dan waktu pemrosesannya 16,63 detik.

Berikut merupakan tabel hasil dari pengujian kesepuluh dokumen diatas, dapat dilihat pada tabel 4.5 dan tabel 4.6 rata – rata dari nilai icf dan presentase kekuatan *stemmer*.

Tabel 4.5 Hasil pengujian dokumen *stemming* dengan Idris dahulu kemudian ECS *Stemmer*

No	Nama Dokumen	Jumlah kata **sebelum** stemming	Algoritma Idris				Algoritma ECS Stemmer			
			Jumlah kata **setelah** stemming	Index Compression (icf)	Waktu Pemrosesan(s)	Presentase kekuatan stemmer	Jumlah kata **setelah** stemming	Index Compression (icf)	Waktu Pemrosesan(s)	Presentase kekuatan stemmer
1	Percy Jackson dan Pencuri Petir	239	209	0,12552	4,37	87%	216	0,09234	1,7	90%

2	The Hunger Games	261	227	0,13027	4,51	86%	229	0,12261	3,14	88%
3	Harry Potter dan Piala Api	323	276	0,14551	6,06	85%	277	0,14241	1,8	85%
4	Dalam Mihrab Cinta	511	415	0,18787	8,75	81%	422	0,17417	2,41	82%
5	Harry Potter dan Batu Bertuah	608	512	0,15789	12,87	84%	533	0,12336	1,71	87%
6	Twilight - Midnight Sun	1023	852	0,16716	20,93	83%	905	0,11535	2,5	88%
7	Percy Jackson dan Lautan Monster	2224	1850	0,16817	38,3	83%	1935	0,12995	5,73	87%
8	If I Stay	3653	3109	0,14892	64,49	85%	3123	0,14509	9,59	85%
9	The Demigod Diaries	4464	3746	0,16094	92,37	84%	3822	0,14382	12,98	85%
10	The Lord of The Rings - Pembawa Sembilan Cincin	5815	4450	0,23474	106	76%	4673	0,19639	16,9	80%
Rata- Rata				0,1627	35,865	83%		0,14368	5,846	86%

Tabel 4.6 Hasil pengujian dokumen *stemming* dengan ECS *Stemmer* dahulu kemudian Idris

No	Nama Dokumen	Jumlah kata **sebelum** stemming	Algoritma Idris				Algoritma ECS Stemmer			
			Jumlah kata **setelah** stemming	Index Compression (icf)	Waktu Pemrosesan(s)	Presentase kekuatan stemmer	Jumlah kata **setelah** stemming	Index Compression (icf)	Waktu Pemrosesan(s)	Presentase kekuatan stemmer
1	Percy Jackson dan Pencuri Petir	239	209	0,12552	4,75	87%	216	0,09234	2,45	90%
2	The Hunger Games	261	227	0,13027	4,72	86%	229	0,12261	3,75	88%
3	Harry Potter dan Piala Api	323	276	0,14551	6,05	85%	277	0,14241	0,74	85%
4	Dalam Mihrab Cinta	511	415	0,18787	7,94	81%	422	0,17417	2,33	82%
5	Harry Potter dan Batu Bertuah	608	512	0,15789	13,03	84%	533	0,12336	1,54	87%
6	Twilight - Midnight Sun	1023	852	0,16716	22,34	83%	905	0,11535	2,42	88%

7	Percy Jackson dan Lautan Monster	2224	1850	0,16817	39.31	83%	1935	0,12995	5,64	87%
8	If I Stay	3653	3109	0,14892	71,19	85%	3123	0,14509	10,7	85%
9	The Demigod Diaries	4464	3746	0,16094	89,57	84%	3822	0,14382	13,17	85%
10	The Lord of The Rings - Pembawa Sembilan Cincin	5815	4450	0,23474	103,91	76%	4673	0,19639	16,62	80%
Rata- Rata				0,1627	35,9444	83%		0,14368	5,936	86%

4.2.2 Skenario Pengujian 2

Pada pengujian ke skenario 2 ini dilakukan untuk mengetahui keakuratan hasil *stemming* oleh masing – masing algoritma baik algoritma Idris dan Algoritma ECS *Stemmer*. Kata – kata yang gagal *stemming* dapat disebabkan oleh skema pemotongan imbuhan yang tidak dapat mencari kata dasar walaupun imbukan sudah dipotong dan sudah dikembalikan jika mengalami salah imbuhan. Berikut ini merupakan contoh kata yang tidak berhasil distemming oleh Algoritma Idris maupun Algoritma ECS *Stemmer* dapat dilihat pada tabel 4.7 dibawah ini.

Tabel 4.7 Contoh kata yang tidak berhasil distemming

No	Kata	Hasil Stemming		Kata dasar yang benar
		Algoritma Idris	Algoritma ECS Stemmer	
1	geletar	geletar	geletar	getar
2	mengantuk	antuk	antuk	kantuk
3	serabut	serabut	rabut	sabut
4	temali	temali	temali	tali
5	kinerja	kinerja	kinerja	kerja
6	mengantuk	antuk	antuk	kantuk
7	mengarang	arang	arang	karang

Kemudian setelah dilakukan analisis kembali pada kata – kata yang telah berhasil distemming tetapi terdapat beberapa kesalahan yang ditemukan karena kata dasar tersebut mengalami *overstemming* atau pemotongan huruf yang seharusnya

tidak dilakukan atau kelebihan memotong huruf sehingga merubah kata dasar yang seharusnya. Berikut contoh pemenggalan imbuhan yang berlebihan atau yang seharusnya tidak dilakukan pemotongan dapat dilihat pada tabel 4.8 dibawah ini.

Tabel 4.8 Kata yang mengalami kesalahan *stemming*

No	Kata	Hasil Stemming		Kata dasar yang benar	Tipe kesalahan
		Algoritma Idris	Algoritma ECS Stemmer		
1	negeri	neger	neger	negeri	*overstemming*
2	buat	bu	bu	buat	kesalahan pemotongan
3	kelas	as	las	kelas	*overstemming*
4	sampai	sampa	sampa	sampai	*overstemming*
5	mengenali	nali	nal	kenal	*overstemming*

Dalam melakukan *stemming* dengan kedua algortima ini yaitu algoritma Idris dan Algoritma ECS *Stemmer* ada beberapa kata yang memang tidak bisa distemming hal ini dikarena kata tersebut merupakan kata dasar yang belum tercantum pada saat pembuatan kamus kata dasar atau kata tersebut bukan kata dasar yang tercantum dalam kamus bahasa Indonesia. Kata – kata seperti diatas menyebabkan *term* tidak dapat distemming dengan benar. Berikut merupakan kata yang tidak terdaftar dalam kamus dapat dilihat pada tabel 4.9 dibawah ini.

Tabel 4.9 Kata yang tidak terdaftar dalam kamus

Kata dasar		
markup	hembus	bongkok
kastil	astronot	akupuntur
negeri	lesat	bolpen
akting	klon	congkel
aksesoris	fesyen	
masing	nampak	

4.3 Kesimpulan

Dari tabel 4.5 dan tabel 4.6 dapat disimpulkan bahwa semakin sedikit kata yang berhasil distemming maka nilai icf aka semakin besar. Nilai icf dari kedua algoritma yaitu algoritma Idris dan algoritma ECS *Stemmer* tidak terlalu jauh berbeda artinya kedua algoritma tersebut dapat diterapkan dengan baik untuk *stem* bahasa Indonesia tetapi nilai icf algoritma ECS *Stemmer* lebih kecil dibandingkan algoritma Idris dapat disimpulkan algoritma ECS *Stemmer* yang paling baik dalam menstemming bahasa Indonesia, selain itu dalam hal persentase kekuatan stemmernya pun algoritma ECS *Stemmer* lebih unggul dari pada algoritma Idris walaupun nilai rata – rata persentasenya tidak terlalu jauh terdapat beda 3%. Untuk setiap dokumen rata – rata terjadi pengurangan kata sebesar 10% sampai 25% setelah dilakukan *stemming* dari jumlah kata awal dokumen sebelum melakukan *stemming.*

Berdasarkan waktu pemrosesan algoritma yang paling cepat dalam menstemming adalah algoritma ECS *Stemmer* perbedaan waktu tersebut terpaut jauh, rata – rata untuk *stemming* pada tabel 4.5 dengan ECS *Stemmer* yaitu 5,846 detik sedangkan algoritma Idris rata – rata *stemming* dokumen sebesar 35,865 detik, pada tabel 4.6 juga dapat dilihat bahwa algoritma ECS *Stemmer* lebih cept dalam menstemming kata dengan rata – rata 5,936 detik sedangkan algortima Idris rata – ratanya 35,94 detik. Jadi dapat ditarik kesimpulan bahwa berdasarkan akurasi dan kecepatan algoritma ECS *Stemmer* lebih baik digunakan untuk menstemming dokumen teks bahasa Indonesia dibandingkan dengan algoritma Idris. Akan tetapi kedua algoritma diatas sama-sama memiliki kekurangan dalam hal gabungan kata yang merupakan gabungan dari dua kata yang menjadi satu seperti kata “prasejarah” merupakan kata yang berasal dari “pra” dan “sejarah”. Kata seperti ini tidak dapat diatasi dengan menggunakan aturan pemotongan imbuhan tetapi perlu dilakukan penyesuai algoritma baru untuk dapat menyelesaikan permasalahan gabungan dua kata yang menjadi satu.