
**RANCANG BANGUN APLIKASI PENCARIAN KEMIRIPAN
DOKUMEN TEKS TUGAS AKHIR MENGGUNAKAN ALGORITMA
RABIN-KARP
(STUDI KASUS : PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS SAINS DAN TEKNOLOGI JAYAPURA)**

Ninny Andriani Mangintiku

Evanita Veronica Manullang

**Program Studi Teknik Informatika
Fakultas Ilmu Komputer Dan Manajemen
Universitas Sains Dan Teknologi Jayapura**

Abstrak - Pesatnya perkembangan teknologi menyebabkan semakin banyaknya informasi yang tersedia, hal tersebut dapat mempermudah seseorang dalam melakukan penjiplakan suatu karya tulis. Perkembangan computer yang meluas dan peran computer dalam mengolah informasi dengan memiliki fasilitas copy-paste (salin-tempel) yang memudahkan dalam pengolahan data sesuai dengan kebutuhan misalnya dalam memenuhi tugas kuliah, membuat paper, membuat tugas akhir dan sebagainya. Hal ini tentunya sangat berpotensi seseorang melakukan penjiplakan dari suatu karya tulis atau sering disebut melakukan plagiat. Plagiarism atau sering disebut plagiat adalah penjiplakan atau pengambilan karya, karangan, pendapat dan sebagainya dari orang lain dan mengakui seolah-olah karangan dan pendapatnya sendiri. Penelitian ini bertujuan untuk membangun sebuah aplikasi pencarian kemiripan dokumen teks pada tugas akhir dengan menggunakan algoritma rabin-karp yang dapat membantu dosen dalam mengecek plagiat atau kemiripan dari penulisan tugas akhir. Aplikasi pencarian kemiripan dokumen teks pada tugas akhir dibangun dengan menggunakan bahasa pemrograman C# dan Mysql sebagai databaseny. Aplikasi yang dihasilkan dapat menampilkan hasil pendeteksian dari isi dokumen yang dicarikan kemiripannya, selain itu aplikasi juga dapat melakukan pencarian kemiripan dalam bentuk format yang berbeda yaitu : .pdf, .docx, dan .txt. dengan adanya aplikasi pencarian kemiripan tugas akhir ini diharapkan dapat membantu para dosen dalam mengecek kesamaan atau kemiripan dari penulisan tugas akhir mahasiswa difakultas ilmu computer dan manajemen.

Kata kunci : Plagiat, Rabin-Karp, Tugas Akhir, Mahasiswa, C#, Mysql.

1. Pendahuluan

1.1 Latar Belakang

Ilmu pengetahuan dan teknologi saat ini berkembang sangat pesat dan semakin maju. Peran komputer saat ini juga semakin meluas tidak hanya alat bantu hitung tetapi juga alat bantu menyelesaikan masalah-masalah yang dihadapi manusia secara cepat, tepat dan efisien. Pesatnya

perkembangan teknologi menyebabkan semakin banyaknya informasi yang tersedia hal tersebut dapat mempermudah seseorang dalam melakukan penjiplakan suatu karya. Pengolahan informasi dalam dunia teknologi yang menggunakan komputer memiliki fasilitas *copy-paste* (salin-tempel) yang memungkinkan pengolahan data sesuai dengan kebutuhan misalnya memenuhi tugas kuliah , membuat jurnal, pembuatan tugas akhir dan sebagainya.

Plagiarisme atau sering disebut sebagai plagiat adalah penjiplakan atau pengambilan karangan, pendapat, dan sebagainya dari orang lain dan menjadikannya seolah-olah karangan dan pendapat sendiri. Di dunia pendidikan dalam melakukan penyusunan Tugas Akhir mahasiswa, sering ditemukan adanya penjiplakan dalam penentuan judul dari Tugas Akhir dimana mahasiswa mengajukan judul kepada dosen berupa judul yang dijiplak dari judul sebelumnya. Dengan cara mengganti metode yang digunakan atau menggunakan dua istila dengan makna yang sama.

Terkadang kemampuan dosen dalam melakukan kontrol dan menyeleksi Tugas Akhir masih terkendala dengan kemampuan daya ingat dari masing-masing dosen atau pembimbing.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka perumusan masalah yang akan dibahas adalah Bagaimana menerapkan algoritma Rabin-Karp untuk membantu dosen dalam mencari kemiripan Tugas Akhir?

1.3 Tujuan Penelitian

Setelah membahas latar belakang, rumusan masalah serta batasan masalah, adapun tujuan dari penelitian yang dilakukan yaitu

1. Untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Teknik Informatika Strata Satu (S1) Fakultas Ilmu Komputer dan Manajemen Universitas Sains dan Teknologi Jayapura.
2. Merancang bangun aplikasi pencarian kemiripan dokumen teks tugas akhir dengan menggunakan Algoritma rabin-karp yang dapat membantu dosen dalam mencari kemiripan dokumen Tugas Akhir.

1.4 Ruang Lingkup Penelitian

Batasan permasalahan pada aplikasi pencarian kemiripan dokumen teks Tugas Akhir menggunakan algoritma rabin-karp, yaitu:

1. Digunakan untuk mencari kemiripan dokumen Tugas Akhir judul dan abstrak.

-
2. Menggunakan bahasa pemrograman *C#*, yang terkoneksi dengan Mysql sebagai basis data.
 3. Mencari kemiripan dengan dokumen yang berformat *.pdf, *.docx, *.txt
 4. Menampilkan nilai presentasi kemiripan dari judul dan abstrak Tugas Akhir.

1.5 Metodologi Penelitian

Untuk mendukung dalam melakukan penelitian dan penyusunan tugas akhir ini, penulis menggunakan beberapa metode dalam pengumpulan data dan juga dalam perencanan sistem. Berikut ini metode-motede yang digunakan yaitu:

1. Metode Kepustakaan

Penulisan ini dimulai dengan studi kepustakaan, yaitu dengan mengumpulkan bahan-bahan referensi baik dari makalah, jurnal, artikel, situs internet yang menyangkut dengan pencarian kemiripan dokumen teks pada Tugas Akhir untuk menunjang pencapaian tujuan Tugas Ahkir yang dibuat.

2. Observasi

Observasi dilakukan untuk melakukan penelitian secara langsung atau melakukan pengamatan terhadap proses yang berjalan dalam pembuatan Tugas Akhir di fakultas FIKOM untuk menunjang pencapaian Tugas Akhir yang akan dibangun.

2. Tinjauan Pustaka

Dian Novian, dkk (2012), dalam bentuk makala dengan judul “Aplikasi Pendeteksian plagiat Pada Karya Ilmiah Menggunakan Metode Rabin Karp” dengan tools yang dipakai yaitu Microsoft Visual C++ dan *SQLite*. Hasil dari penelitian ini ialah dapat mendeteksi plagiat pada karya ilmiah dan memperoleh pelaporan hasil plagiat.

Dodoli Aria Putra, ddk (2015), dalam bentuk jurnal dengan judul “Implementasi Algoritma Rabin Karp Untuk Membantu Pendeteksian Plagiat Pada Karya Ilmiah” dengan menggunakan PHP dan *MySQL*. Hasil dari penelitian ini ialah dapat mendeteksi plagiat dan membandingkan kecepatan pendeteksian dari karya ilmiah.

Edwin Wijaya (2014 / 2015), dalam bentuk makala dengan judul “Pencocokan Pola Majemuk Dengan Lgoritma Rabin – Karp” dengan menggunakan JAVA sebagai toolsnya, pendeteksian yang dilakukan yaitu mencocokkan pola dalam teks majemuk dan memperoleh hasil pencocokan pola majemuk dari suatu karya

Moh Faud Uddin (2016), dalam bentuk artikel skripsi dengan judul “Sistem Deteksi Kemiripan Judul Skripsi Program Studi Teknik Informatika Menggunakan Algoritma Rabin – Karp” dengan tools yang digunakan ialah Delphi 7. Hasil dari penelitian ini ialah pengguna dapat melihat tingkat kemiripan dari judul skripsi yang akan diajukan oleh mahasiswa.

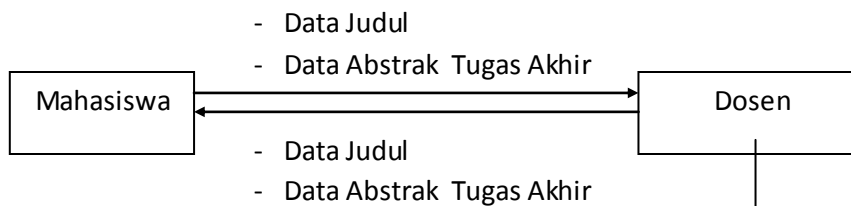
Kadek Versi Yana Yoga (2012) , dalam bentuk artikel dengan judul “Pengembangan Aplikasi Pendeteksian Plagiasme Pada Dokumen Teks Menggunakan Algoritma Rabin – Karp” hasil dari penelitian ini ialah dapat mendeteksi plagiat pada dokumen teks berbahasa Indonesia dan mendeteksi kata yang bersinonim.

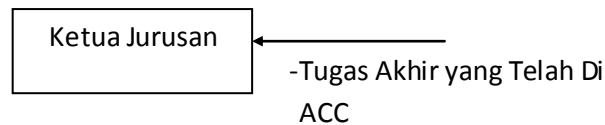
3. Analisa Dan Rancangan

3.1 Analisa Sistem

1. Sistem Berjalan

Adapun sistem yang sedang berjalan di Jurusan Teknik Informatika Fakultas Ilmu Komputer dan Manajemen Universitas Sains dan Teknologi Jayapura sebagai berikut:





Gambar 1. Sistem Berjalan

2. Studi Kasus

Dalam sebuah kasus akan dilakukan proses pencarian kemiripan antara dua judul dan dua abstrak / intisari dari masing – masing tugas akhir, proses perhitungan manual dari studi kasus tersebut adalah sebagai berikut :

Penulis :	Qomariyah AL Qibtiyah ; Agustus; 2014
Judul :	Sistem pakar untuk mendiagnosa penyakit pada organ reproduksi wanita dengan menggunakan metode Dempster Shafer.
Abstrak :	<p>Penyakit pada sistem organ reproduksi wanita adalah penyakit yang perlu dipertimbangkan, karena kadang kala fungsi – fungsi sistem organ reproduksi tidak dapat berjalan sebagaimana mestinya, sedangkan pengetahuan tentang gejala dan penyakit pada sistem organ reproduksi itu sendiri masih banyak yang belum diketahui oleh kaum wanita dan hanya menganggap gejala umum biasa, namun tidak menutup kemungkinan gejala yang timbul adalah gejala penyakit yang dapat menyebabkan kematian.</p> <p>Tujuan penelitian ini membuat aplikasi sistem pakar (Expert System) berbasis web sehingga dapat membantu masyarakat khususnya wanita untuk mengetahui jenis penyakit pada sistem organ reproduksi wanita dengan menggunakan Metode <i>Certainty Faktor</i> (CF), metode yang digunakan untuk membangun sebuah aplikasi sistem pakar yang dikembangkan menggunakan bahasa pemograman <i>PHP</i> dan <i>MySQL</i> sebagai penyimpanannya.</p> <p>Hasil dari penelitian menunjukkan tingkat kebenaran dari kemungkinan penyakit pada sistem organ reproduksi wanita terhadap kemungkinan gangguan–gangguan penyakit yang diderita serta informasi pengobatan selanjutnya.</p> <p>Kata kunci : Sistem Pakar, Certainty Factor, Sistem</p>

	reproduksi wanita, PHP- My SQL.
Penulis :	Jeni Ati Tandi Bulu ; Januari: 2014
Judul :	Sistem pakar berbasis web untuk mendiagnosa gangguan penyakit pada sistem organ reproduksi wanita menggunakan metode Certainty Factor
Abstrak :	<p>Teknologi komputer dewasa ini terus menunjang kehidupan manusia sehingga komputer bukan lagi menjadi hal yang asing bagi sebagian besar orang Sistem pakar (<i>Artificial Intelegence</i>).Sistem pakar adalah perangkat lunak yang di desain khusus berdasarkan <i>Artificial Intelegence</i> yang berfungsi untuk merekam dan menduplikasikan kemampuan pakar.Penelitian ini mengadopsi pengetahuan pakar penyakit yang berpengalaman sebagai contoh, membangun sistem pakar untuk mendiagnosa penyakit pada Organ Reproduksi wanita dengan metode <i>Dempster Shafer</i>.Sistem pakar ini berguna untuk mengetahui informasi mengenai gejala – gejala penyakit yang sering kali diabaikan oleh para wanita.</p> <p>Tujuan dari penelitian ini adalah memberikan suatu informasi kepada masyarakat yang dikhususkan pada wanita tentang penyakit pada Organ Reproduksi dan membuat suatu perangkat lunak dapat mengidentifikasi penyakit dengan menggunakan rekayasa sistem pakar, sedangkan metode yang digunakan adalah metode <i>Dempster Shafer</i> dimana metode ini merupakan teori matematika dari <i>evidence</i>, teori tersebut dapat memberikan sebuah cara untuk mnggabungkan <i>evidence</i> dari beberapa sumber dan mendatangkan / memberikan tingkat kepercayaan dimana kebutuhan probabilitas yang akan dibutuhkan untuk setiap pertanyaan dari keinginan.</p> <p>Hasil dari penelitian ini adalah sistem pakar dapat memberkan suatu informasi kepada masyarakat tentang penyakit yang diderita sehingga ditanggulangi sejak dini.</p> <p>Kata kunci : Sistem Pakar, Metode Dempster Shafer, Reproduksi Wanita, Artificial Intelegence.</p>

Tabel 1. Hasil Parsing K-gram Judul Tugas Akhir

No	Parsing File 1	Nilai Hashing	Parsing File 2	Nilai Hashing
1.	Siste		Siste	
2.	Istem		Istem	
3.	stemp		stemp	
4.	tempa		tempa	
5.	empak		empak	
88.	Actor

Untuk mendapatkan hasil atau nilai dari hasing maka digunakan rumus sebagai berikut :

Keterangan:

K-gram : 5

Modulo : 101

Bentuk perhitungan untuk mencari nilai hash adalah sebagai berikut :

$$H = C_1 * b^{(K-1)} + C_2 * b^{(K-2)} + \dots + C_{(K-1)} * b^{(K)} + c_k \dots$$

Contoh perhitungan penulis mengambil akan melakukan perhitungan dengan parsing "siste", "istem", "stemp", "tempa", "empak" untuk parsing K-gram 1 dan parsing 2.

$$\begin{aligned} \text{"Siste"} &= s * 10^{(5-1)} + i * 10^{(5-2)} + s * 10^{(5-3)} + t * 10^{(5-4)} + e * 10^{(5-5)} \\ &= 115 * 10^4 + 105 * 10^3 + 115 * 10^2 + 116 * 10^1 + 101 * 10^0 \\ &= 1150000 + 105000 + 11500 + 1160 + 101 \\ &= 1267761 \end{aligned}$$

$$\text{Modulo} = 1267761 \text{ mod } 101 = 9$$

$$\begin{aligned} \text{"Istem"} &= i * 10^{5-1} + s * 10^{5-2} + t * 10^{5-3} + e * 10^{5-4} + m * 10^{5-5} \\ &= 105 * 10^4 + 115 * 10^3 + 116 * 10^2 + 101 * 10^1 + 109 * 10^0 \\ &= 1050000 + 115000 + 11600 + 1010 + 109 \\ &= 177719 \end{aligned}$$

$$\text{Modulo} = 177719 \text{ mod } 101 = 59$$

$$\begin{aligned} \text{"stemp"} &= s * 10^{5-1} + t * 10^{5-2} + e * 10^{5-3} + m * 10^{5-4} + p * 10^{5-5} \\ &= 115 * 10^4 + 116 * 10^3 + 101 * 10^2 + 109 * 10^1 + 112 * 10^0 \\ &= 1150000 + 116000 + 10100 + 1090 + 112 \\ &= 1277302 \end{aligned}$$

$$\text{Modulo} = 1277302 \text{ mod } 101 = 56$$

Table 2. Nilai hasing dari parsing K-gram

No	Parsing	Nilai	Nilai	Parsing	Nilai	Nilai
----	---------	-------	-------	---------	-------	-------

	File 1	Hashing	Modulo	File 2	Hasing	Modulo
1.	Siste	1267761	9	siste	1267761	9
2.	Istem	1177719	59	istem	1177719	59
3.	stemp	1277302	56	stemp	1277302	56
4.	tempa	1273117	12	tempa	1273117	12
5.	empak	1131277	77	empak	1131277	77
6.	Mpaka	1212867	59	mpaka	1212867	59

Tabel 3. Perhitungan untuk nilai abstrak :

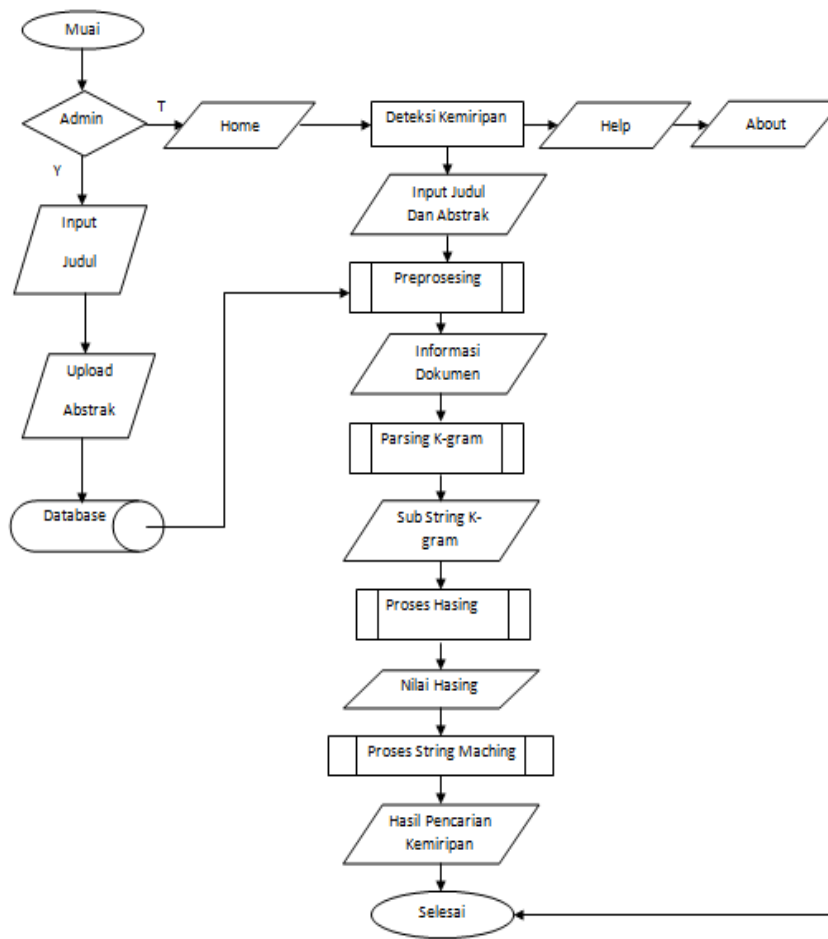
No	Parsing File 1	Nilai Hashing	Nilai Modulo	Parsing File 2	Nilai Hasing	Nilai Modulo
1.	intis			intis		
2.	ntisa			ntisa		
3.	tisar			tisar		
4.	isari			isari		
5.	saris			sarit		
6.	arisa			arite		
20.	organ

Tabel 4. Perhitungan untuk nilai abstrak

No	Parsing File 1	Nilai Hashing	Nilai Modulo	Parsing File 2	Nilai Hasing	Nilai Modulo
1.	intis	1172765	54	intis	1172765	54
2.	ntisa	1227747	92	ntisa	1227747	92
3.	tisar	1277584	35	tisar	1277584	35
4.	isari	1175945	2	isari	1175945	2
5.	saris	1259565	95	sarit	1259566	96
6.	arisa	1095747	99	arite	1095761	12
20.	organ

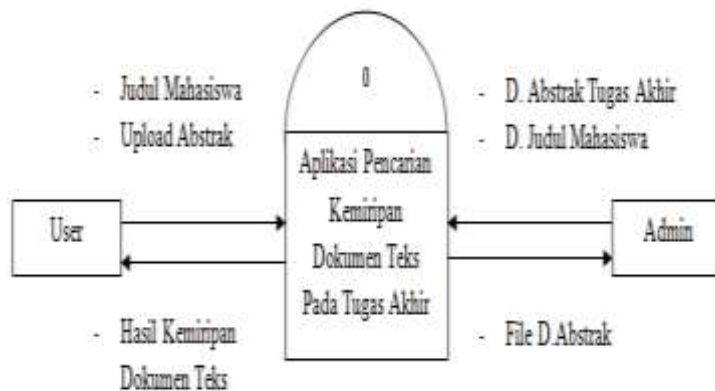
3.2 Perancangan

1. Flowchart



Gambar 2. Flowchart

2. Diagram Konteks

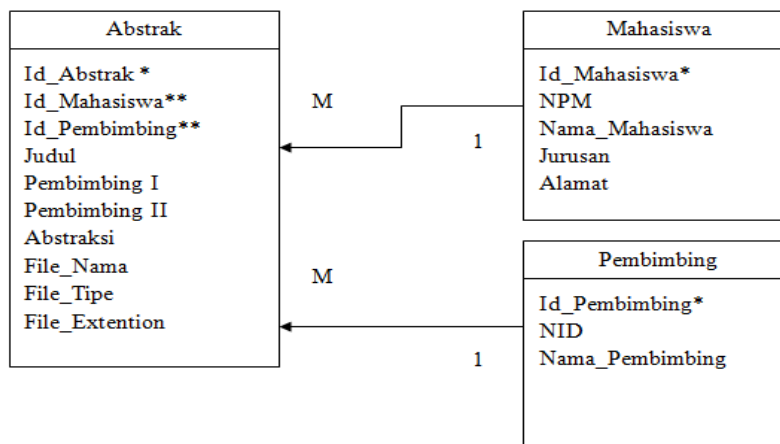


3. Diagram Overview Level 0



Gambar 4. DFD Level 0

4. Skema Tabel



Gambar 5. Skema table aplikasi yang dibangun

4. Hasil Dan Pembahasan

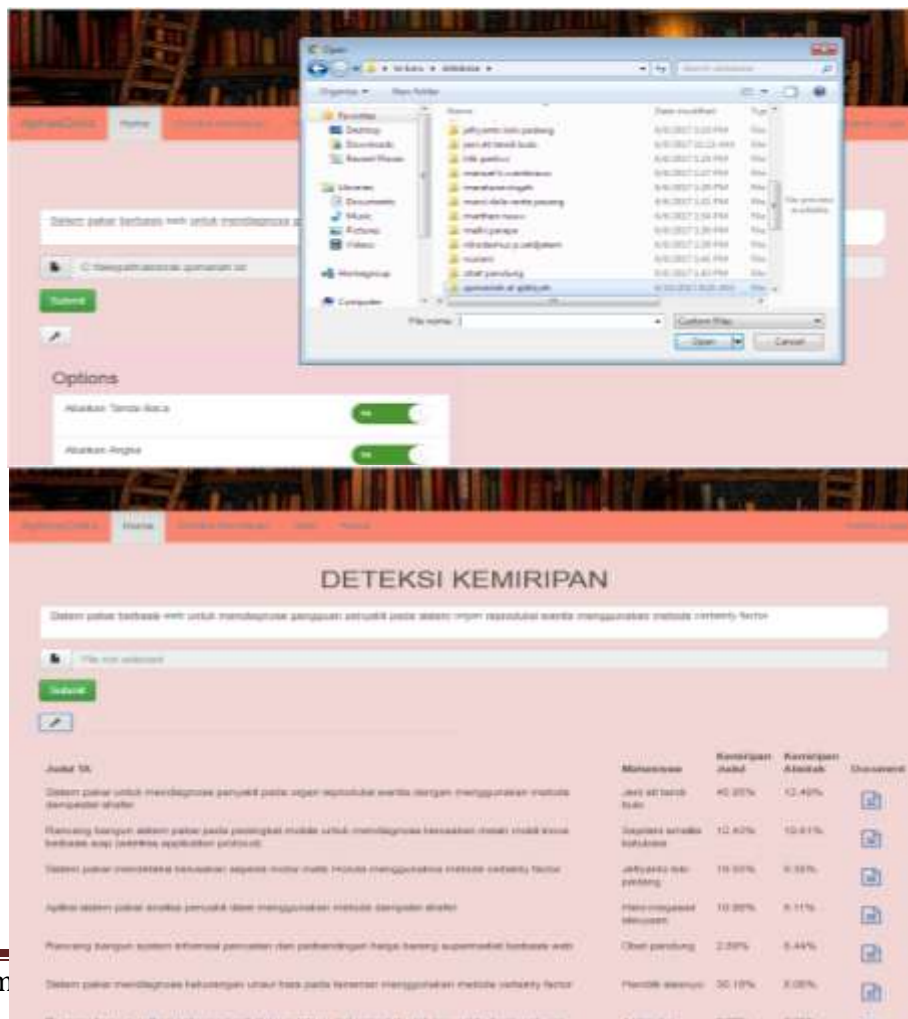
4.1 Pengujian Halaman User

Halaman user hanya menampilkan informasi dan bentuk file yang akan di uji oleh aplikasi pencarian kemiripan dokumen teks pada tugas akhir di program studi teknik informatika, fakultas FIKOM



4.2 Halaman Deteksi kemiripan

Halaman deteksi adalah halaman yang akan melakukan analisa, dimana file yang akan dilakukan analisa diupload lalu di lakukan analisa.



bstrak)

Gambar 8. Halaman Deteksi Kemiripan (hasil presentasi kemiripan)



kakan pencocokan)
utkan), data abstrak

1. Pada halaman deteksi plagiat jika judul tugas akhir dan abstrak tugas akhir tidak diisi, maka sistem akan menolak untuk melakukan penyimpanan data kedalam sistem dan akan menampilkan pesan kesalahan. Pada halaman deteksi plagiat juga memiliki dua tombol yaitu tombol “submit” dan tombol “options”. Tombol “submit” berfungsi untuk melakukan analisa untuk mengetahui tingkat kemiripan dari judul dan abstrak yang telah diinputkan dan diupload sedangkan tombol “option” berfungsi untuk memberikkan settingan file yang akan di uji yaitu apakah file yang di uji akan “mengabaikan tanda baca”, “mengabaikan angka”, dan mengabaikan “” .
2. Jika judul tugas akhir diisi dan abstrak tugas akhir juga diisi, maka sistem akan menyimpan data kedalam database dan akan melakukan proses analisa untuk mengetahui tingkat kemiripan.
3. Setelah hasil pencarian kemiripan dokumen dilakukan maka user dapat melihat dokumen yang diupload dengan menekan gambar dokumen yang berada di halaman deteksi kemiripan.

5. Penutup

5.1 Kesimpulan

1. Aplikasi pencarian kemiripan dokumen teks dapat melakukan pencarian kemiripan dokumen teks dengan format yang dipakai untuk pencarian kemiripan dokumen ialah dokumen yang berformat *.txt, *.doc, *.pdf proses pencarian kemiripan dapat dilakukan dengan cara melakukan input dan upload data yang akan di cari kemiripannya setelah itu aplikasi akan melakukan proses inisialisasi, proses analisa hingga menampilkan presentase dari tingkat kemiripan dokumen.
2. Aplikasi yang dirancang ini dapat membantu pihak dosen dalam melakukan pengawasan terhadap mahasiswa dalam melakukan penyusunan Tugas Akhir sehingga dosen dapat membantu dalam mengurangi kemiripan judul Tugas Akhir mahasiswa hingga file data tugas akhir mahasiswa.

5.2 Saran

1. Melihat dari bentuk tampilan aplikasi yang dirancang dan dibuat oleh penulis dalam penyusunan tugas akhir ini penulis masih terdapat kekurangan dari segi tampilan aplikasi sehingga kedepannya untuk pengembangan aplikasi ini agar lebih di perbaiki lagi dari segi tampilan untuk dapat lebih menarik perhatian bagi user dalam menggunakan aplikasi ini.
2. Melihat dari tingkat durasi atau lamanya waktu dalam menganalisa dokumen teks yang diupload dan di inputkan oleh pengguna/user. Proses menginisialisasi file, proses convert file agar bisa dibaca oleh system dan dianalisa hingga proses memunculkan hasil analisa yang membutuhkan waktu lama maka penulis harapkan kepada mahasiswa yang akan melakukan pengembangan terhadap aplikasi ini agar membuat perbaikan waktu atau durasi dalam melakukan analisa hingga memunculkan hasil presentasi kemiripan.
3. Melihat dari proses algoritma yang dipakai yang berhasil memunculkan hasil analisa algoritma hingga hasil pencarian kemiripan dokumen teks maka penulis harapkan agar kedepannya jika dilakukan pengembangan untuk aplikasi ini agar melakukan pergantian algoritma yang dipakai sehingga dapat menyimpulkan bagaimana proses berjalannya algoritma lain yang dipakai apakah berhasil menampilkan presentasi kemiripan ataukah tidak.

Daftar Pustaka

Fathansyah, 2012, *Basis Data*, Edisi Revisi, Informatika, Bandung

Hidayatullah, P., Kawistara, K.J., 2015, *Pemograman Web*, Informatika, Bandung. Kamus Besar Bahasa Indonesia, 2001.

Kendall, E. Kendall., Kendall, E.J., 2010, *Analisa dan Perancangan Sistem*, Indeks, Jakarta.

Departemen Pendidikan Bahasa Indonesia, 2008, "*Kamus Besar Bahasa Indonesia*" PT. Gramedia Pustaka Utama. Jakarta.

Novian Dian, 2012, "*Aplikasi Pendeteksian Plagiat Pada Karya Ilmiah Menggunakan Metode Rabin-Karp*" E-Book.pdf (Diakses tanggal 12 Mei, Jam 12.00 am)

Putra Aria Dodoli, 2015, "*Implementasi Algoritma Rabin-Karp dalam Membantu Pendeteksian Plagiat Pada Karya Ilmiah*" E-Book.pdf (Diakses tanggal 12 Mei, Jam 12.00 am)

Uddin Faud Moh, 2016, "*Sistem Deteksi Kemiripan Judul Skripsi Program Studi Teknik Informatika Menggunakan Algoritma Rabin-Karp*", E-Book.pdf (Diakses tanggal 12 Mei, Jam 12.00 am)

Wijaya Edwin, 2015, "*Pencocokan Pola Majemuk Dengan Algoritma Rabin-Karp*", E-Book.pdf (Diakses tanggal 12 Mei, Jam 12.00 am)

Yakub, 2012, "*Pengantar Sistem Informasi*", Graha Ilmu, Yogyakarta.

Yoga Yana Versi Kadek, 2012, "*Pengembangan Aplikasi Pendeteksian Plagiarisme Pada Dokumen Teks Menggunakan Algoritma Rabin-Karp*", E-Book.pdf (Diakses tanggal 12 Mei, Jam 12.00 am)