

**IMPLEMENTASI ALGORITMA RABIN-KARP PADA PENENTUAN  
TEMATIK SKRIPSI**

**SKRIPSI**



disusun oleh

**Reno Diandika Hermawan**

**15.11.8878**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2019**

**IMPLEMENTASI ALGORITMA RABIN-KARP PADA PENENTUAN  
TEMATIK SKRIPSI**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika

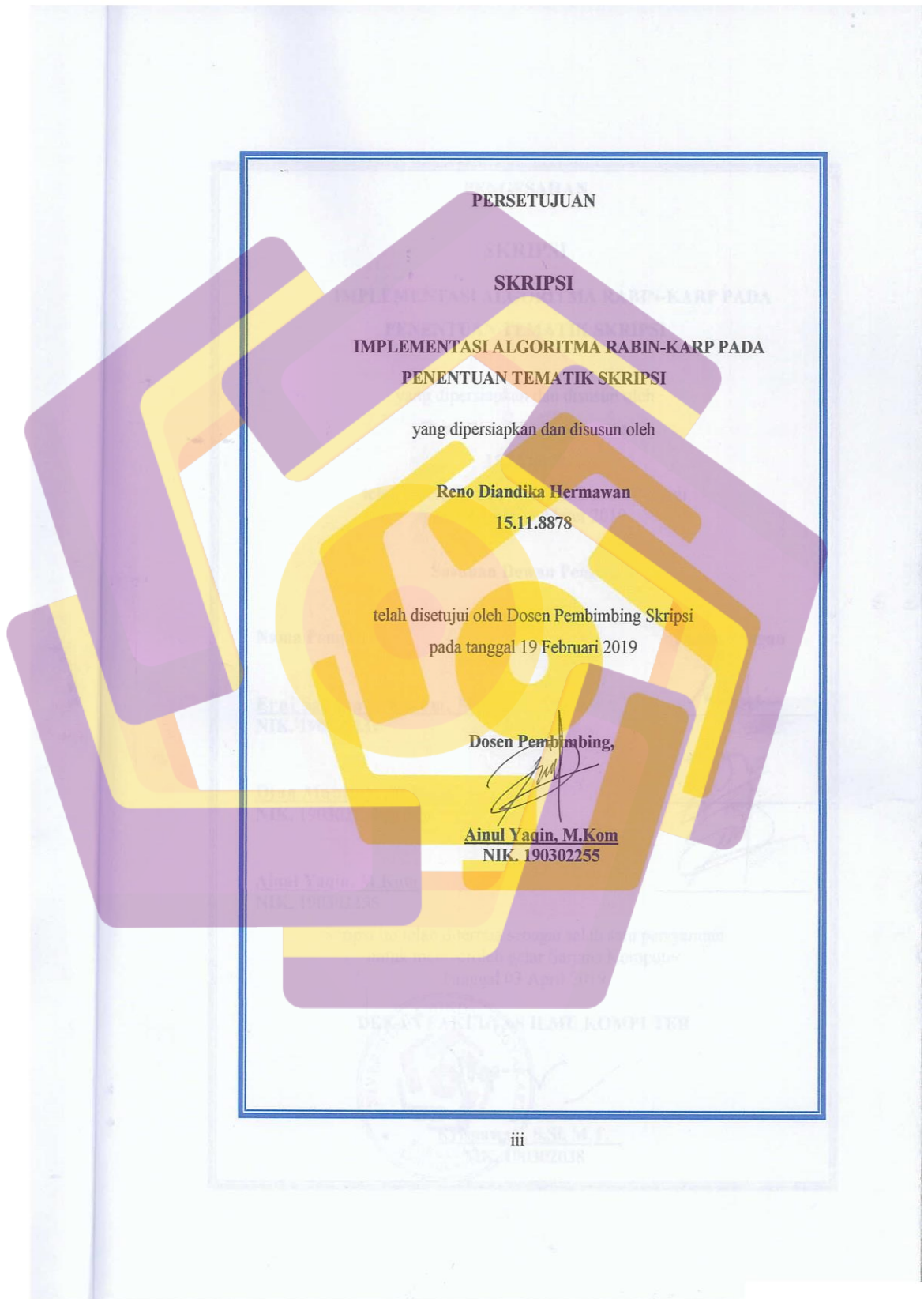


disusun oleh

**Reno Diandika Hermawan**

**15.11.8878**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2019**



**PERSETUJUAN**

**SKRIPSI**

**IMPLEMENTASI ALGORITMA RABIN-KARP PADA  
PENENTUAN TEMATIK SKRIPSI**

yang dipersiapkan dan disusun oleh

**Reno Diandika Hermawan**

**15.11.8878**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 19 Februari 2019

**Dosen Pembimbing,**

**Ainal Yaqin, M.Kom**  
**NIK. 190302255**

<b>PENGESAHAN</b>	
<b>SKRIPSI</b>	
<b>IMPLEMENTASI ALGORITMA RABIN-KARP PADA</b>	
<b>PENENTUAN TEMATIK SKRIPSI</b>	
yang dipersiapkan dan disusun oleh	
<b>Reno Diandika Hermawan</b>	
<b>15.11.8878</b>	
telah dipertahankan di depan Dewan Penguji	
pada tanggal 19 Maret 2019	
<b>Susunan Dewan Penguji</b>	
<b>Nama Penguji</b>	<b>Tanda Tangan</b>
<u><b>Erni Seniwati, S.Kom, M.Cs</b></u> NIK. 190302231	
<u><b>Dina Maulina, M.Kom</b></u> NIK. 190302250	
<u><b>Ainul Yaqin, M.Kom</b></u> NIK. 190302255	
Skripsi ini telah diterima sebagai salah satu persyaratan	
untuk memperoleh gelar Sarjana Komputer	
Tanggal 03 April 2019	
<b>DEKAN FAKULTAS ILMU KOMPUTER</b>	
	
	
<u><b>Krisnawati, S.Si, M.T.</b></u> NIK. 190302038	

MOTTO

**PERNYATAAN**

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 09 April 2019



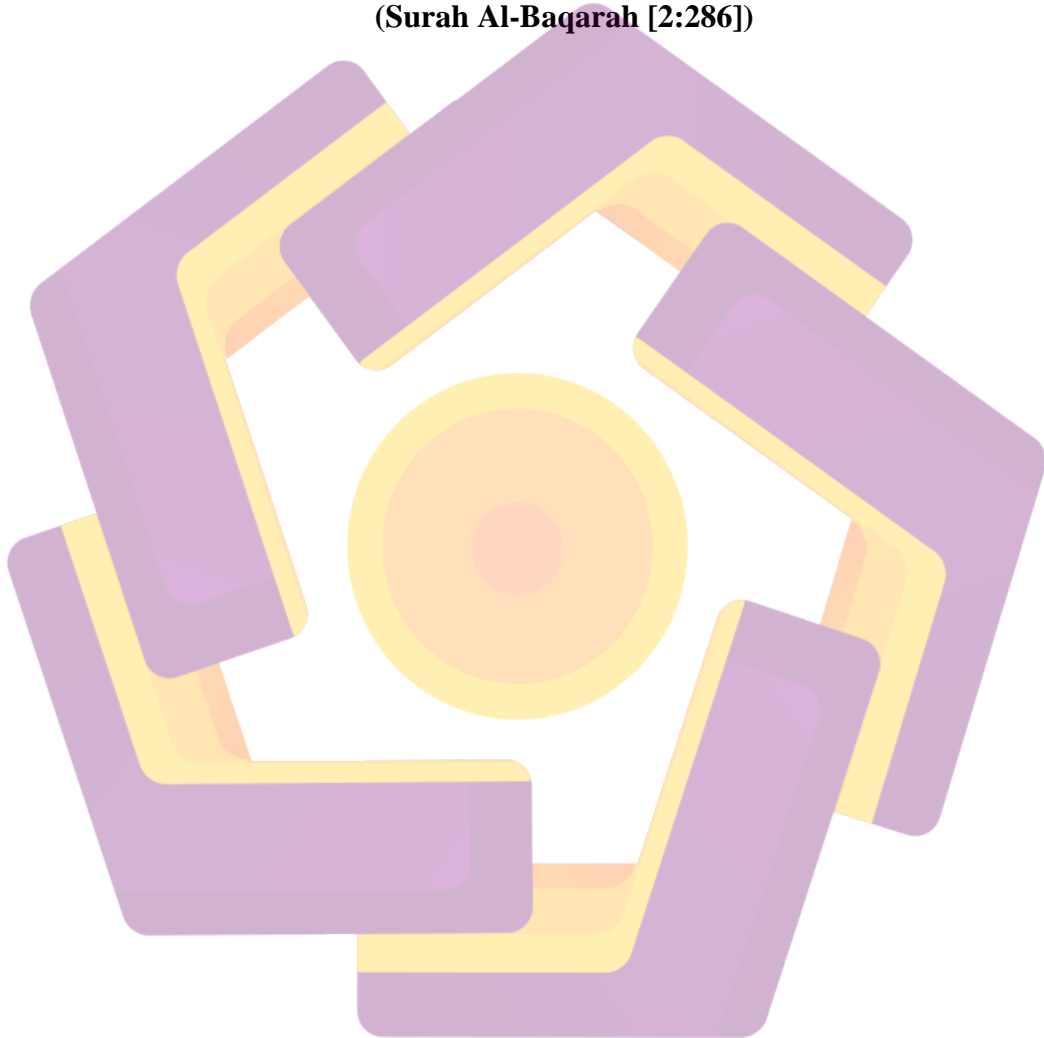
Reno Diandika Hermawan

NIM. 15.11.8878

## MOTTO

“Allah tidak membebani seseorang melainkan sesuai dengan dengan  
kesanggupannya.”

**(Surah Al-Baqarah [2:286])**



## PERSEMBAHAN

*Alhamdulillahirobbil'alamiin*, akhirnya Penulis dapat menyelesaikan skripsi ini dengan baik. Pada kesempatan ini penulis tak lupa mengucapkan terimakasih kepada:

1. Allah SWT atas kehendak-Nya skripsi ini bisa selesai dengan baik.
2. Semua anggota keluarga terima kasih karena telah memberikan kesempatan dan waktu.
3. Bapak Ainul Yaqin, M.Kom selaku dosen pembimbing, terima kasih telah memberikan saran dalam penyusunan skripsi ini.
4. Teman-teman 15 S1-Informatika 06 terima kasih telah menjadi teman seperjuangan dalam menuntut ilmu selama ini.
5. Serta semua pihak yang telah membantu tersusunnya skripsi ini yang tidak dapat saya sebutkan satu persatu.

## KATA PENGANTAR

Alhamdulillah Rabbil Alamin, terima kasih atas kehendak-Nya sehingga penulis masih diberikan kesempatan untuk menyelesaikan penulisan penelitian yang berjudul "Implementasi Algoritma Rabin-Karp pada Penentuan Tematik Skripsi" ini tanpa adanya halangan yang berarti. Penelitian ini memuat tentang bagaimana algoritma Rabin-Karp yang sering digunakan untuk mendeteksi plagiarisme diimplementasikan kedalam pengklasifikasian tema skripsi berdasarkan abstrak.

Penulis menyadari bahwa penulisan penelitian ini masih banyak terdapat kekurangan, dipersilahkan untuk mengambil manfaatnya, bila ada salahnya itu karena kesalahan dari penulis sendiri. Terimakasih kepada seluruh pihak yang telah memberikan dukungan dalam penyelesaian penelitian ini, semoga apa yang telah diberikan dapat bernilai sebagai amalan baik.

Yogyakarta, 09 April 2019

Reno Diandika Hermawan



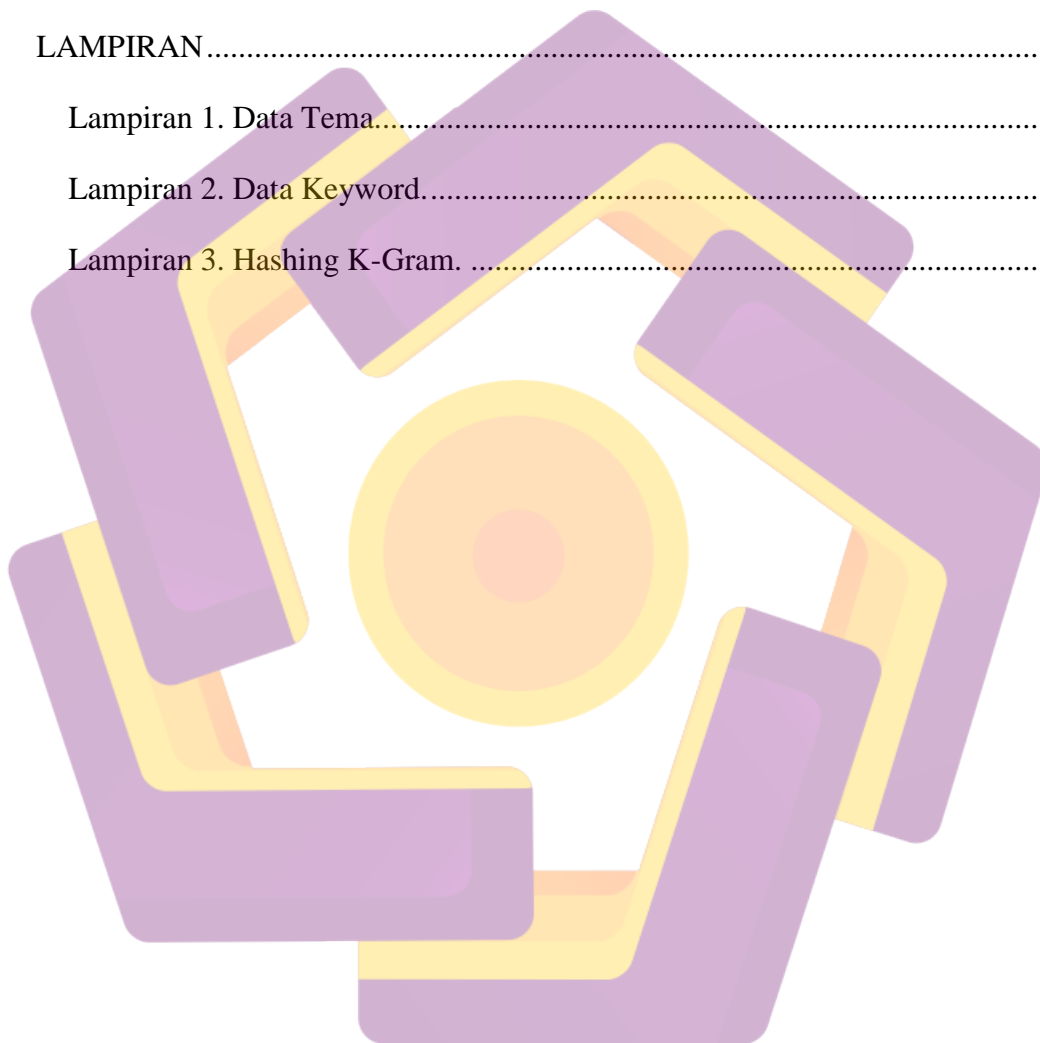
## DAFTAR ISI

COVER .....	i
LEMBAR JUDUL .....	ii
PERSETUJUAN .....	iii
PENGESAHAN .....	iv
PERNYATAAN .....	v
MOTTO .....	vi
PERSEMBAHAN .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
INTISARI .....	xv
<i>ABSTRACT</i> .....	xvi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	2
1.4 Maksud dan Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3

1.6 Metode Penelitian.....	3
1.6.1 Metode Penelitian.....	3
1.6.2 Metode Pengumpulan Data .....	3
1.6.3 Metode Analisis .....	3
1.6.4 Metode Perancangan .....	4
1.6.5 Metode Pengembangan .....	4
1.6.6 Metode Testing.....	4
1.7 Sistematika Penulisan .....	4
<b>BAB II LANDASAN TEORI.....</b>	<b>6</b>
2.1 Kajian Pustaka.....	6
2.2 Dasar Teori.....	9
2.2.1 Tematik .....	9
2.2.2 <i>Natural Language Processing</i> .....	9
2.2.3 Ekstraksi Dokumen ( <i>Text Preprocessing</i> ) .....	9
2.2.4 <i>K-gram</i> .....	10
2.2.5 <i>Hashing</i> .....	10
2.2.6 Algoritma Rabin-Karp .....	11
2.2.7 <i>Dice's Similarity Coeficients</i> .....	11
2.2.8 <i>Data Flow Diagram (DFD)</i> .....	12
2.2.9 <i>Entity Relationship Diagram (ERD)</i> .....	13
2.2.10 <i>Website</i> .....	14
2.2.11 <i>Hypertext Preprocessor (PHP)</i> .....	14
2.2.12 MySQL.....	14
<b>BAB III METODE PENELITIAN.....</b>	<b>15</b>
3.1 Alur Penelitian .....	15
3.2 Identifikasi Masalah .....	16
3.3 Pengumpulan Data .....	16
3.4 Pembuatan Flowchart Rabin-Karp.....	17

3.5	Perhitungan Manual Algoritma Rabin-Karp .....	19
3.5.1	Abstrak Skripsi dan Keyword .....	19
3.5.2	<i>Text Processing</i> .....	19
3.5.3	Proses Rabin-Karp .....	21
3.6	Perancangan Proses .....	29
3.6.1	<i>Data Flow Diagram (DFD)</i> .....	29
3.7	Perancangan Tabel .....	31
3.7.1	<i>Entity Relathionship Diagram (ERD)</i> .....	31
3.7.2	Relasi Tabel Tematik Skripsi (db_tematik) .....	32
3.7.3	Perancangan Tabel .....	33
<b>BAB IV IMPLEMENTASI DAN PEMBAHASAN .....</b>		<b>35</b>
4.1	Implementasi Database .....	35
4.2	Implementasi Antarmuka .....	35
1.	Menu Login .....	35
2.	Menu Tematik Skripsi .....	36
3.	Menu Tambah Tema .....	37
4.	Menu Hapus Tema .....	37
5.	Menu Tambah Keyword .....	38
6.	Menu Hapus Keyword .....	39
7.	Menu Help .....	39
4.3	Implementasi Source Code .....	40
4.3.1	<i>Case folding &amp; Tokenizing</i> .....	40
4.3.2	<i>Stopword Removal</i> .....	40
4.3.3	<i>Parsing K-Gram</i> .....	40
4.3.4	<i>Hashing K-Gram</i> .....	41
4.3.5	Menjumlah <i>Hash Value</i> .....	42
4.3.6	Menjumlah Banyaknya <i>K-Gram</i> .....	42
4.3.7	<i>Text Similarity</i> .....	43
4.4	Pembahasan Pengujian .....	43

BAB V PENUTUP.....	45
5.1 Kesimpulan .....	45
5.2 Saran.....	45
DAFTAR PUSTAKA .....	46
LAMPIRAN.....	48
Lampiran 1. Data Tema.....	48
Lampiran 2. Data Keyword.....	48
Lampiran 3. Hashing K-Gram. ....	50



## DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Sebelumnya.....	7
Tabel 2.2 <i>Data Flow Diagram</i> (DFD) .....	12
Tabel 2.3 <i>Entity Relationship Diagram</i> (ERD).....	13
Tabel 3.1 Abstrak Skripsi dan Keyword.....	19
Tabel 3.2 <i>Text Processing</i> .....	20
Tabel 3.3 Kata Abstrak Skripsi dan Keyword .....	22
Tabel 3.4 Nilai <i>hash</i> sama.....	24
Tabel 3.5 <i>Similarity</i> .....	26
Tabel 3.6 <i>Multi_login</i> .....	33
Tabel 3.7 Skripsi .....	33
Tabel 3.8 Tema.....	33
Tabel 3.9 Keyword.....	33
Tabel 3.10 <i>Kata_hubung</i> .....	34
Tabel 5.1 Hasil Pengujian.....	43

## DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	16
Gambar 3.2 Flowchart Rabin-Karp.....	18
Gambar 3.3 Diagram Konteks Tematik Skripsi.....	29
Gambar 3.4 DFD Tematik Skripsi Level 1 .....	30
Gambar 3.5 DFD Tematik Skripsi Level 2 Proses Tematik Skripsi .....	31
Gambar 3.6 ERD Tematik Skripsi .....	32
Gambar 3.7 Relasi Tabel Tematik Skripsi .....	32
Gambar 4.1 Db_tematik.....	35
Gambar 4.2 Menu Login.....	36
Gambar 4.3 Menu Tematik Skripsi.....	36
Gambar 4.4 Menu Tambah Tema .....	37
Gambar 4.5 Menu Hapus Tema .....	38
Gambar 4.6 Menu Tambah <i>Keyword</i> .....	38
Gambar 4.7 Menu Hapus <i>Keyword</i> .....	39
Gambar 4.8 Menu <i>Help</i> .....	39
Gambar 4.9 <i>Case folding &amp; Tokenizing</i> .....	40
Gambar 4.10 <i>Stopword Removal</i> .....	40
Gambar 4.11 <i>Parsing K-Gram</i> .....	41
Gambar 4.12 <i>Hashing K-Gram</i> .....	41
Gambar 4.13 Menjumlah <i>Hash Value</i> .....	42
Gambar 4.14 Menjumlah <i>Banyaknya K-Gram</i> .....	42
Gambar 4.15 <i>Text Similarity</i> .....	43
Gambar 4.16 Diagram Hasil Pengujian <i>Confusion Matrix</i> .....	44

## INTISARI

Skripsi / karya ilmiah adalah komposisi ilmu yang menyajikan fakta dan ditulis sesuai dengan metodologi penulisan yang baik dan benar. Tetapi untuk mengetahui apa tema yang dibahas harus membaca abstrak esai. Seiring dengan perkembangan teknologi informasi yang menawarkan kemudahan bagi tenaga manusia dalam penyederhanaan karyanya. Kesulitan dalam menentukan tema dalam tesis dapat dibantu dengan aplikasi yang mampu mengatur tema berdasarkan abstrak. Untuk membuat aplikasi penentuan tematik skripsi ini dapat menggunakan algoritma pencocokan string, salah satu algoritma pencocokan string adalah algoritma Rabin-Karp dengan melakukan pencocokan string berdasarkan nilai hash pada teks dan nilai hash dalam suatu pola.

Algoritma Rabin-Karp sering digunakan untuk mendeteksi dokumen teks plagiarisme, membandingkan nilai hash keseluruhan k-gram yang terkandung dalam semua dokumen teks, kemudian memutuskan tingkat kesamaan. Maka dapat disimpulkan bahwa Rabin-Karp mampu merangkai teks secara keseluruhan, dari sini peneliti tertarik untuk membuktikan bagaimana jika algoritma Rabin-Karp diimplementasikan dalam bidang klasifikasi di mana dalam penelitian ini digunakan untuk pencocokan string setiap kata atau kata kunci mengacu pada tema esai atau abstrak tesis Indonesia.

Aplikasi ini dibangun menggunakan bahasa pemrograman web PHP dan database MySQL. Matriks kebingungan digunakan untuk menentukan akurasi, diperoleh 88% dan 12% untuk tingkat kesalahan klasifikasi. Tematik Skripsi kurang bermanfaat jika berdiri sendiri, ini bisa lebih bermanfaat jika dikombinasikan dengan sistem pendukung keputusan yang terkait dengan akademisi, di mana dapat memanfaatkan hasil tesis tematik untuk menghasilkan keputusan yang lebih baik atau lebih akurat.

**Kata Kunci:** Skripsi, Abstrak, Tematik, String Matching, Rabin-Karp, Hash.

## ABSTRACT

*Thesis / scientific work is a science composition that presents facts and is written according to a good and correct writing methodology. But to find out what the theme discussed should read the abstract of the essay. Along with the development of information technology that offers convenience for human labor in simplification of his work. Difficulty in determining the theme in the thesis can be aided with applications that are able to set themes based on abstract. To make the application of thesis thematic determination can use string matching algorithm, one of string matching algorithm is Rabin-Karp algorithm by doing string matching based on the hash value on text and hash value in a pattern.*

*The Rabin-Karp algorithm is often used to detect plagiarism text document, comparing the overall hash value of k-gram contained in all text documents, then deciding the degree of similarity. Then it can be concluded that Rabin-Karp is able to string a text as a whole, from this the researcher is interested in proving what if the Rabin-Karp algorithm is implemented in the classification field wherein in this study it is used to string matching each word or keyword refers to the theme of an Indonesian essay or thesis abstract.*

*The application is built using web programming language PHP and MySQL database. The confusion matrix is used to determine accuracy, obtained 88% and 12% for misclassification rates. Thematic Thesis lacks benefits if it stands alone, this can be more useful if combined with a decision support system related to academics, where it can utilize the results of thematic thesis to produce better or more accurate decisions.*

**Keyword:** *Thesis, Abstract, Thematic, String Matching, Rabin-Karp, Hash.*