

**ANALISIS TINGKAT KESAMAAN DOKUMEN TEKS  
MENGUNAKAN RABIN KARP**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**DESMAR ZUHDIL AMRI**

**17.11.0965**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**ANALISIS TINGKAT KESAMAAN DOKUMEN TEKS  
MENGUNAKAN RABIN KARP**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**DESMAR ZUHDIL AMRI**

**17.11.0965**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS TINGKAT KESAMAAN DOKUMEN TEKS**

**MENGGUNAKAN RABIN KARP**

yang disusun dan diajukan oleh

**Desmar Zuhdil Amri**

**17.11.0965**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 21 November 2022

**Dosen Pembimbing**



**Hastari Utama, M.Cs**

**NIK. 190302230**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS TINGKAT KESAMAAN DOKUMEN TEKS**

**MENGGUNAKAN RABIN KARP**

yang disusun dan diajukan oleh

**Desmar Zuhdil Amri**

**17.11.0965**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 21 November 2022

**Susunan Dewan Penguji**

**Nama Penguji**

**Wahid Miftahul Ashari, S.Kom., M.T**  
**NIK : 190302452**

**Robert Marco, M.T.**  
**NIK : 190302228**

**Hastari Utama, M.Cs**  
**NIK : 190302230**

**Tanda Tangan**



Handwritten signatures of the examiners, including one that appears to read 'Hastari'.

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 21 November 2022

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, S.Kom., M.Kom.**  
**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Desmar Zuhdil Amri  
NIM : 17.11.0965

Menyatakan bahwa Skripsi dengan judul berikut:

### **ANALISIS TINGKAT KESAMAAN DOKUMEN TEKS MENGUNAKAN RABIN KARP**

Dosen Pembimbing : Hastari Utama, M.Cs

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 November 2022

Yang Menyatakan,



Desmar Zuhdil Amri

## KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah Swt. atas ridho-Nya saya dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang saya ajukan adalah “Analisis Tingkat Kesamaan Dokumen Teks Menggunakan Rabin Karp”

Skripsi ini diajukan untuk memenuhi syarat kelulusan mata kuliah Skripsi di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta. Namun, karya ini tidak akan selesai tanpa orang-orang di sekeliling saya yang mendukung dan membantu.

Terima kasih saya sampaikan kepada:

Hastari Utama, M.Cs., selaku Dosen pembimbing yang telah memberikan bimbingan dan berbagai pengalaman kepada penulis dan Semua pihak yang telah membantu dan tidak dapat disebutkan satu per satu.

Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah Swt. dan akhirnya saya menyadari bahwa skripsi ini masih jauh dari kata sempurna, karena keterbatasan ilmu yang saya miliki. Untuk itu saya dengan kerendahan hati mengharapkan saran dan kritik yang sifatnya membangun dari semua pihak demi membangun laporan penelitian ini.

Yogyakarta, 21 November 2022

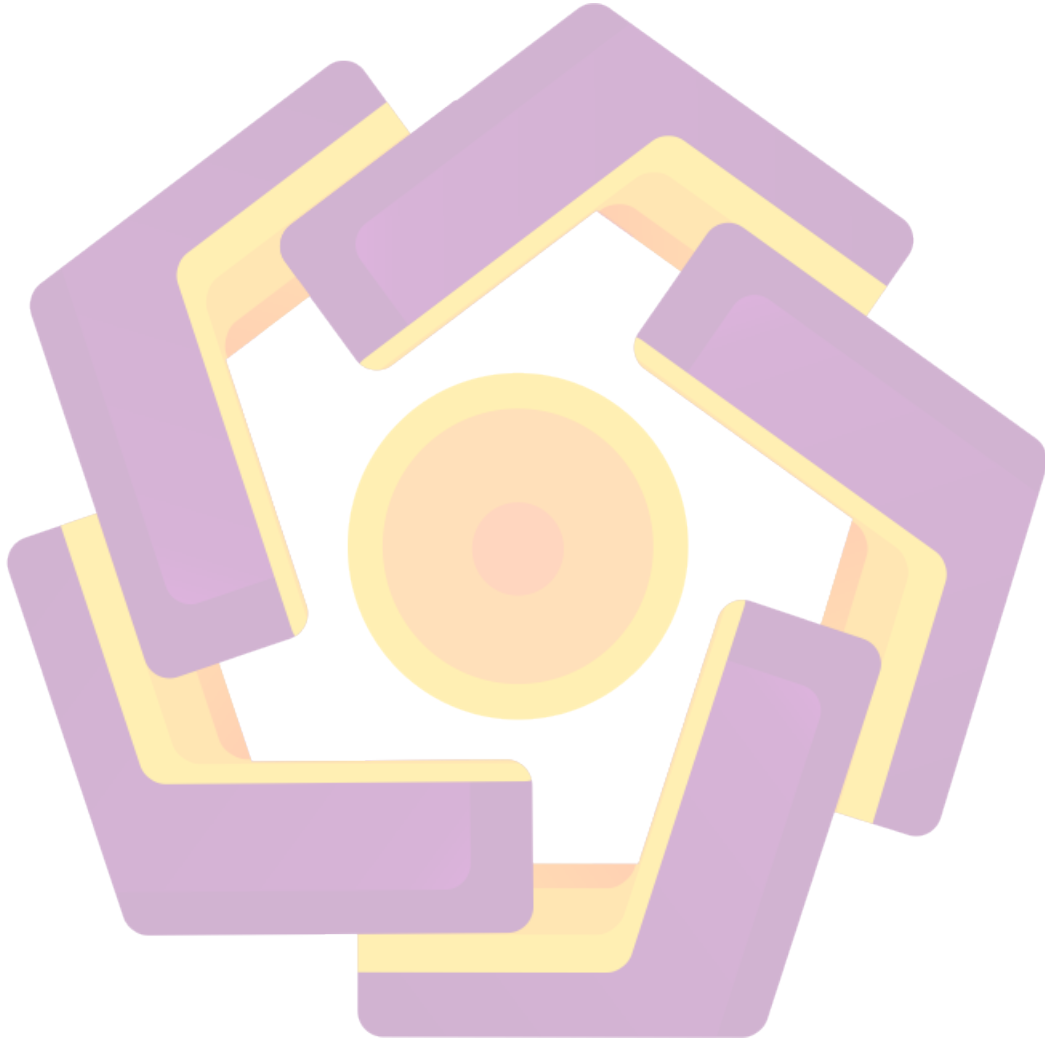


Desmar Zuhdil Amri

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>2</b>
<b>HALAMAN PERSETUJUAN .....</b>	<b>3</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>4</b>
<b>HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....</b>	<b>5</b>
<b>KATA PENGANTAR.....</b>	<b>6</b>
<b>DAFTAR ISI.....</b>	<b>7</b>
<b>DAFTAR TABEL .....</b>	<b>9</b>
<b>DAFTAR GAMBAR.....</b>	<b>10</b>
<b>INTISARI .....</b>	<b>11</b>
<b>ABSTRACT .....</b>	<b>12</b>
<b>BAB I PENDAHULUAN.....</b>	<b>13</b>
<b>1.1 Latar Belakang .....</b>	<b>13</b>
<b>1.2 Rumusan Masalah.....</b>	<b>14</b>
<b>1.3 Batasan Masalah .....</b>	<b>14</b>
<b>1.4 Tujuan penelitian .....</b>	<b>14</b>
<b>1.5 Manfaat Penelitian .....</b>	<b>14</b>
<b>BAB II LANDASAN TEORI .....</b>	<b>15</b>
<b>2.1 Kajian Pustaka .....</b>	<b>15</b>
<b>2.2 Dasar Teori .....</b>	<b>16</b>
<b>BAB III METODE PENELITIAN .....</b>	<b>20</b>
<b>3.1 Tahapan Penelitian .....</b>	<b>20</b>
<b>3.2 Studi Literatur.....</b>	<b>20</b>
<b>3.3 Identifikasi Permasalahan.....</b>	<b>21</b>
<b>3.4 Alur Algoritma Rabin Karp.....</b>	<b>21</b>
<b>3.5 Perhitungan Manual .....</b>	<b>22</b>
<b>BAB IV RANCANGAN.....</b>	<b>31</b>

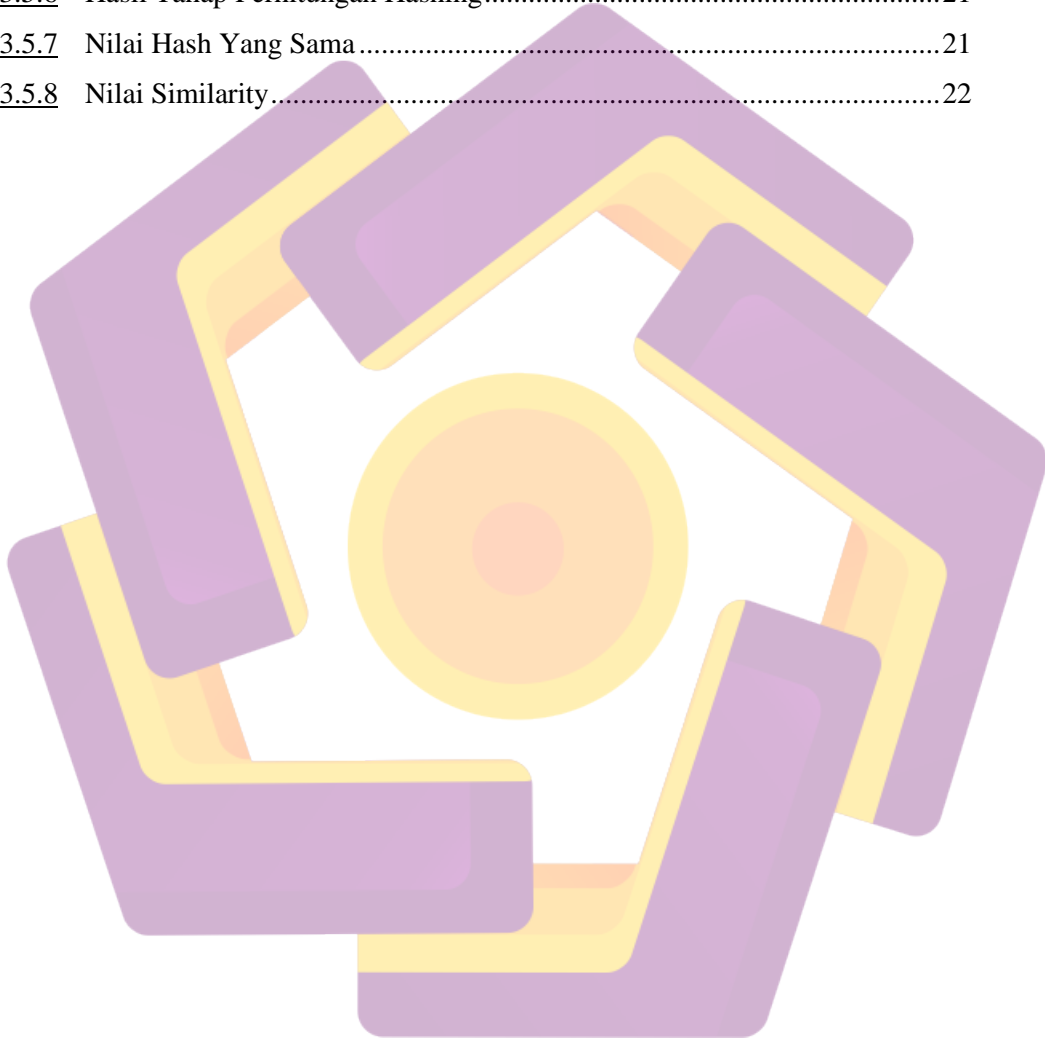
<b>4.1</b>	<b>Halaman Pertama</b> .....	<b>31</b>
<b>4.2</b>	<b>Halaman Hasil</b> .....	<b>31</b>
<b>4.3</b>	<b>Uji Coba Sistem</b> .....	<b>32</b>
<b>BAB V KESIMPULAN DAN SARAN</b> .....		<b>34</b>
<b>5.1</b>	<b>Kesimpulan</b> .....	<b>34</b>
<b>5.2</b>	<b>Saran</b> .....	<b>34</b>
<b>DAFTAR PUSTAKA</b> .....		<b>35</b>





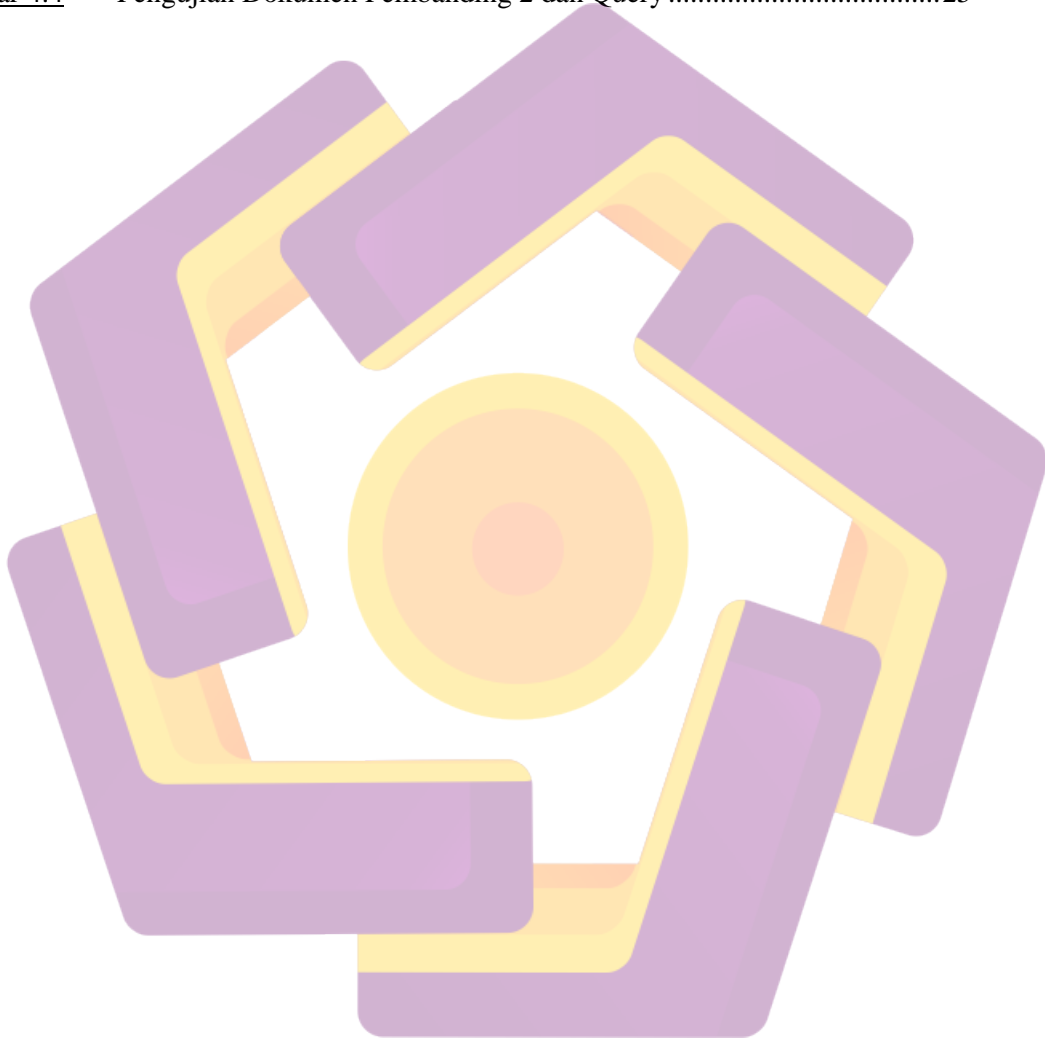
## DAFTAR TABEL

<u>Tabel 3.5.1</u>	<u>Dokumen Teks</u> .....	15
<u>Tabel 3.5.2</u>	<u>Hasil Preprocessing</u> .....	15
<u>Tabel 3.5.3</u>	<u>Tokenizing k-gram</u> .....	16
<u>Tabel 3.5.4</u>	<u>Karakter ASCII</u> .....	17
<u>Tabel 3.5.5</u>	<u>Nilai Kode ASCII</u> .....	17
<u>Tabel 3.5.6</u>	<u>Hasil Tahap Perhitungan Hashing</u> .....	21
<u>Tabel 3.5.7</u>	<u>Nilai Hash Yang Sama</u> .....	21
<u>Tabel 3.5.8</u>	<u>Nilai Similarity</u> .....	22



## DAFTAR GAMBAR

<u>Gambar 3.1</u>	<u>Alur Penelitian</u> .....	12
<u>Gambar 3.2</u>	<u>Alur Algoritma Rabin Karp</u> .....	13
<u>Gambar 4.1</u>	<u>Gambar Uji 1</u> .....	23
<u>Gambar 4.2</u>	<u>Gambar Uji 2</u> .....	23
<u>Gambar 4.3</u>	<u>Pengujian Dokumen Pembanding 1 dan Query</u> .....	24
<u>Gambar 4.4</u>	<u>Pengujian Dokumen Pembanding 2 dan Query</u> .....	25



## INTISARI

Plagiarisme didefinisikan sebagai mengambil atau mencuri beberapa karya dan mengklaim menjadi miliknya sendiri. Dengan kata lain, plagiarisme adalah suatu aksi merugikan seperti halnya tindakan penipuan. Seperti mengambil karya orang lain setelah itu mengaku bahwa karya tersebut adalah hasil kerja originalnya. Kasus plagiat juga bisa mempengaruhi dunia akademik, dikarenakan menghambat perkembangan kreativitas untuk membuat karya ilmiah menggunakan ide yang berasal dari pikiran sendiri. Dengan adanya kasus plagiat maka harus ada system yang dapat mendeteksi plagiarisme. Pada dasarnya sistem yang dibuat untuk mendeteksi plagiat rata-rata menerapkan pencocokan string di dalam sebuah teks. Algoritma sederhana yang sering digunakan adalah rabin karp, dikarenakan rabin karp dapat mencocokkan beberapa pola sekaligus yang terdapat didalam string. Algoritma rabin karp tidak melakukan proses lama untuk mendeteksi teks, karena rabin karp menyaring karakter yang tidak cocok dan setelah itu melakukan perbandingan dalam teks. Rata-rata hasil kesamaan teks yang dapat dideteksi cukup tinggi dikarenakan semakin kecil nilai k-gramnya maka semakin banyak pola string yang terdeteksi.

***Kata kunci*** : algoritma Rabin Karp, kesamaan dokumen teks, plagiarisme, similarity

## ABSTRACT

Plagiarism is defined as taking or stealing some works and claiming to be his own. In other words, plagiarism is an act of detriment as well as fraud. It's like taking someone else's work and then admitting that the work is their original work. Cases of plagiarism can also affect the academic world, because it hinders the development of creativity to make scientific work using ideas that come from one's own mind. With plagiarism cases, there must be a system that can detect plagiarism. Basically a system created to detect plagiarism, on average, applies string matching in a text. The simple algorithm that is often used is rabin karp, because rabin karp can match several patterns contained in a string at once. The rabin karp algorithm does not take a long time to detect text, because it filters out unmatched characters and then performs a comparison in the text. The average result of text similarity that can be detected is quite high because the smaller the k-gram value, the more string patterns are detected.

***Keywords :***

Algoritm rabin karp, similarity document text, plagiarism, similarity

