

SKRIPSI
APLIKASI PENDETEKSI KEMIRIPAN DOKUMEN
BERBASIS WEB MENGGUNAKAN
ALGORITMA RABIN-KARP



Oleh :
Muhammad Renaldi Prayoga
NIM : 205410096

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA
2022

SKRIPSI

**APLIKASI PENDETEKSI KEMIRIPAN DOKUMEN
BERBASIS WEB MENGGUNAKAN
ALGORITMA RABIN-KARP**

Diajukan sebagai salah satu syarat untuk menyelesaikan studi



Disusun Oleh

MUHAMMAD RENALDI PRAYOGA

NIM : 205410096

**PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA**

2022

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 26 Januari 2022



Muhammad Renaldi Prayoga
NIM: 205410096

HALAMAN PERSEMBAHAN

Alhamdulillah rabbi 'alamin atas rahmat dan hidayah yang telah diberikan Allah SWT sehingga dapat terwujudnya skripsi ini dengan penuh perjuangan dan pengorbanan. Semoga skripsi ini dapat menjadi sesuatu yang bermanfaat dan berguna serta menjadi kebanggaan untuk keluarga tercinta. Karya yang sederhana ini penulis persembahkan untuk :

1. Allah SWT. Yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini tepat waktu dan dengan baik.
2. Kedua orang tua saya Bapak Bambang Hermanto, dan Ibu Sri Wahyuni yang selalu mendukung dan memberikan semangat saya sehingga dapat menyelesaikan skripsi ini.
3. Apresiasi sebesar-besarnya kepada diri saya sendiri karena mampu melewati salah satu tangga ujian kehidupan yang akan mengantarkan ke masa depan yang lebih baik untuk kedepannya.
4. Adik saya satu-satunya Dewi Rara Angraini yang menjadi semangat saya untuk menyelesaikan skripsi ini.

5. Ibu Maria Mediatrix Sebatubun S.Kom., M.Eng. Selaku Dosen Pembimbing saya untuk strata 1 yang tiada henti memberikan bimbingan, dorongan, saran, kritik, dan motivasi kepada penulis
6. Bapak M. Agung Nugroho S.Kom., M.Kom. selaku dosen penguji yang telah memberikan saran dan kritik selama proses penyusunan skripsi ini.
7. Teman-teman seperjuangan saya sekaligus teman merantau saya dari Pontianak Rifqi, Ridha, dan Tegar yang mendukung sekaligus memberikan semangat selama berlangsungnya pengerjaan skripsi ini.
8. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan bantuan baik berupa moril maupun materiil, sehingga penulis dapat menyelesaikan skripsi ini.

KATA PENGANTAR

Alhamdulillah, puji syukur kami panjatkan kehadirat Allah SWT, yang telah memberikan rahmat dan hidayah-Nya. Shalawat serta salam kami curahkan kepada Nabi Besar Muhammad SAW, keluarga, sahabat, serta umatnya hingga akhir zaman. Sehingga pada kesempatan ini kami berhasil menyelesaikan penyusunan skripsi dengan berjudul “Aplikasi Pendeteksi Kemiripan Dokumen Berbasis Web Menggunakan Algoritma Rabin-Karp” tepat pada waktunya.

Penulis menyadari bahwa selama proses penyusunan skripsi ini telah mendapatkan banyak bimbingan, saran, kritik, doa dan motivasi dari berbagai macam pihak. Oleh karena itu, penulis mengucapkan banyak terima kasih kepada :

1. Orang tua yang selalu memberikan nasihat dan dukungan selama berjalannya proses penyusunan Skripsi.
2. Bapak Ir. Totok Suprawoto, M.M., M.T. selaku Ketua Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, S.T., M.T. selaku ketua Program Studi Informatika.
4. Ibu Maria Mediatrix Sebatubun S.Kom., M.Eng. selaku Dosen Pembimbing yang tiada henti memberikan kritik dan saran selama pengerjaan skripsi ini.
5. Bapak M. Agung Nugroho S.Kom., M.Kom. selaku dosen penguji yang telah memberikan saran dan kritik selama proses penyusunan skripsi ini.
6. Bapak dan ibu dosen yang telah membagi ilmunya selama penulis masih menempuh perkuliahan di Universitas Teknologi Digital Indonesia.
7. Seluruh staff pengajar dan administrasi Universitas Teknologi Digital Indonesia.
8. Teman-teman seperjuangan saya yang telah memberikan semangat dan motivasi selama penyusunan skripsi ini.

9. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan bantuan baik berupa moril maupun materiil, sehingga penulis dapat menyelesaikan skripsi ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih jauh dari sempurna, oleh karena itu, kritik dan saran dari pembaca yang membangun sangat penulis harapkan untuk menjadi pelajaran di kemudian hari. Penulis berharap semoga skripsi ini dapat memberikan manfaat bagi kita semua. Terima kasih

Pontianak, Januari 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN SKRIPSI.....	v
HALAMAN PERSEMBAHAN.....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABLE.....	xiii
DAFTAR LAMPIRAN.....	xiv
INTISARI.....	xv
ABSTRACT.....	xvi
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Ruang Lingkup.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI.....	4
2.1 Tinjauan Pustaka.....	4
2.2 Dasar Teori.....	6
2.2.1 Kemiripan.....	6
2.2.2 Algoritma Rabin-Karp.....	6
2.2.3 Hashing.....	7
2.2.4 Rolling Hash.....	8
2.2.5 K-Gram.....	8
2.2.6 Text Preprocessing.....	9

2.2.7 Persamaan Dice's Similarity Coefficient.....	10
BAB 3 METODE PENELITIAN.....	11
3.1 Kebutuhan Bahan atau Data.....	11
3.1.1 Kebutuhan Input.....	11
3.1.2 Kebutuhan Proses.....	11
3.1.3 Kebutuhan Output.....	12
3.2 Peralatan.....	12
3.2.1 Kebutuhan Perangkat Keras.....	12
3.2.2 Kebutuhan Perangkat Lunak.....	12
3.3 Metode Pengujian dan Pengumpulan Data.....	13
3.3.1 Metode Literatur.....	13
3.3.2 Metode Pengembangan Aplikasi.....	13
3.4 Analisis dan Rancangan Sistem.....	15
3.4.1 FlowChart.....	16
3.4.2 Rancangan Interface Halaman.....	21
BAB 4 IMPLEMENTASI DAN PEMBAHASAN.....	23
4.1 Implementasi dan Uji Coba Sistem.....	23
4.1.1 Tampilan Halaman Utama.....	23
4.1.2 Tampilan Hasil Pengecekan.....	24
4.2 Pembahasan.....	25
4.2.1 Pengujian.....	26
4.2.2 Hasil Pengujian.....	26
4.2.3 Analisis Hasil dan Pembahasan.....	31
BAB 5 PENUTUP.....	33
5.1 Kesimpulan.....	33
5.2 Saran.....	33
DAFTAR PUSTAKA.....	35
LAMPIRAN.....	37

DAFTAR GAMBAR

Gambar 2.1: Rolling Hash.....	8
Gambar 3.1: Model Waterfall.....	14
Gambar 3.2: Flowchart aplikasi.....	16
Gambar 3.3: Flowchart Preprocessing.....	18
Gambar 3.4: Flowchart Bagian Proses.....	20
Gambar 3.5: Rancangan Tampilan Halaman Website.....	21
Gambar 4.1: Tampilan Halaman Utama.....	24
Gambar 4.2: Tampilan Sebelum Pengecekan Kemiripan.....	24
Gambar 4.3: Tampilan Hasil Uji Coba Sistem.....	25
Gambar 4.4: Grafik Hasil Pengujian dengan nilai K-gram 4.....	28
Gambar 4.5: Grafik Hasil Pengujian dengan nilai K-gram 5.....	29
Gambar 4.6: Grafik Hasil Pengujian dengan nilai K-gram 6.....	31
Gambar 4.7: Grafik Perbandingan Antar K-Gram.....	32

DAFTAR TABLE

Table 4.1: Hasil Pengujian dengan K-Gram Bernilai 4.....	27
Table 4.2: Hasil Pengujian dengan K-Gram Bernilai 5.....	28
Table 4.3: Hasil Pengujian dengan K-Gram Bernilai 6.....	30

DAFTAR LAMPIRAN

Lampiran 1 Listing Program.....	37
Lampiran 2 Cara Menjalankan Program.....	48
Lampiran 3 Kriteria Kelulusan Ujian Sidang/Pendadaran.....	50
Lampiran 4 Catatan Pendadaran Penguji.....	51
Lampiran 5 Catatan Pendadaran Pembimbing.....	52
Lampiran 6 Keputusan Hasil Ujian Pendadaran.....	53
Lampiran 7 Surat Keterangan Persetujuan Publikasi.....	54

INTISARI

Mudahnya mencari informasi melalui internet ini memberikan kesempatan untuk sebagian orang dalam melakukan plagiat. Plagiat sering terjadi di mana saja salah satunya di lingkungan akademik, tindak plagiat mudah dilakukan dengan cara copy-paste. Dengan pengetahuan yang cukup mengenai tindak plagiat dan teknologi informasi yang maju saat ini hal tersebut dapat dicegah.

Dalam menyelesaikan masalah tersebut maka diperlukannya sebuah aplikasi yang dapat melakukan pengecekan pada dokumen yang satu dengan dokumen yang lainnya. Maka dari itu diperlukannya suatu algoritma yang dapat melakukan pencarian pola pada kedua dokumen tersebut. Salah satunya adalah menggunakan algoritma rabin-karp yang mencari pola pada substring dengan menggunakan hash.

Maka dari itu dibuatkan lah suatu aplikasi berbasis desktop yang digunakan untuk melakukan mengecek tingkat kemiripan pada kedua dokumen dengan menerapkan *algoritma rabin-karp*, dengan teknik *hash* yang digunakan adalah *rolling hash*.

Kata Kunci : *Algoritma Rabin-Karp, Hash, Rolling Hash*

ABSTRACT

The ease of finding information through the internet provides an opportunity for some people to do plagiarism. Plagiarism often occurs anywhere, one of which is in the academic environment, plagiarism is easy to do by copy-pasting. With sufficient knowledge about plagiarism and advanced information technology, this can be prevented.

In solving this problem, we need an application that can check one document with another. Therefore we need an algorithm that can perform pattern searches on the two documents. One of them is using the rabin-karp algorithm which looks for patterns in substrings using hashes.

Therefore, a desktop-based application was created that is used to check the level of similarity in the two documents by applying the rabin-karp algorithm, with the hash technique used is rolling hash.

Keywords : Rabin-Karp Algorithm, Hash, Rolling Hash