



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

**PERBANDINGAN RABIN-KARP ALGORITHM DAN JARO
WINKLER DISTANCE ALGORITHM DALAM PENDETEKSI
PLAGIARISME PADA DOKUMEN TEKS**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar

Sarjana Komputer (S.Kom.)



Brinardi Leonardo

12110110029

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN KOMUNIKASI
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2016**

HALAMAN PENGESAHAN SKRIPSI

HALAMAN PENGESAHAN SKRIPSI

PERBANDINGAN RABIN-KARP ALGORITHM DAN JARO WINKLER DISTANCE ALGORITHM DALAM PENDETEKSI PLAGIARISME PADA DOKUMEN TEKS

Oleh

Nama : Brinardi Leonardo
NIM : 12110110029
Program Studi : Teknik Informatika
Fakultas : Teknologi Informasi dan Komunikasi

Tangerang, 22 Agustus 2016

Ketua Sidang

Dosen Penguji



Dr. Ir. P. M. Winarno, M.Kom.



Marcel Bonar Kristanda, S.Kom., M.Sc.

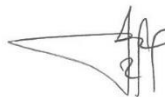
Dosen Pembimbing



Seng Hansun S.Si., M.Cs.

Mengetahui,

Ketua Program Studi
Teknik Informatika



Maria Irima Prasetyowati, S.Kom., M.T.

PERNYATAAN TIDAK MELAKUKAN PLAGIAT

Dengan ini saya,

Nama : Brinardi Leonardo
NIM : 12110110029
Program Studi : Teknik Informatika
Fakultas : Teknologi Informasi dan Komunikasi

Menyatakan bahwa skripsi yang berjudul **“Perbandingan Rabin-Karp Algorithm dan Jaro Winkler Distance Algorithm dalam Pendeteksi Plagiarisme pada Dokumen Teks”** ini adalah karya ilmiah saya sendiri, bukan plagiat dari karya ilmiah yang ditulis oleh orang lain atau lembaga lain, dan semua karya ilmiah orang lain atau lembaga lain yang dirujuk dalam skripsi ini telah disebutkan sumber kutipannya serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah Skripsi yang telah saya tempuh.

Tangerang, 22 Agustus 2016

Brinardi Leonardo

HALAMAN PERSEMBAHAN/MOTO



*Keberhasilan, kesuksesan, dan kebahagiaan
Tidak bisa datang secara seketika
Semua membutuhkan sebuah proses, pengalaman bahkan sebuah kegagalan
Dan
Percayalah Semua sudah dituliskan seperti cerita oleh Sang Pencipta
Jadi teruslah berSyukur apapun hasilnya itu.*

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa berkat keagungan dan kemuliaan dari Nya dalam memberikan kelimpahan pengetahuan kepada penulis sehingga penulis dapat menyelesaikan laporan skripsi ini dengan baik.

Disadari bahwa selesainya penulisan laporan skripsi ini tidak lepas mendapatkan dukungan dan bantuan dari berbagai pihak yang memberikan ilmu, masukan, doa, dan motivasi kepada penulis. Untuk itu izinkan penulis mengucapkan banyak terima kasih kepada:

1. Dr. Ninok Leksono, Rektor Universitas Multimedia Nusantara, yang memberi inspirasi bagi penulis untuk berprestasi,
2. Kanisius Karyono, S.T., M.T., Dekan Fakultas Teknologi Informasi dan Komunikasi Universitas Multimedia Nusantara,
3. Maria Irminda Prasetiyowati, S.Kom., M.T., Ketua Program Studi Teknik Informatika, yang menerima penulis dengan baik untuk berkonsultasi,
4. Seng Hansun, S.Si., M.Cs. selaku wakil kepala program studi Teknik Informatika Universitas Multimedia Nusantara dan juga sebagai pembimbing dalam pembuatan laporan skripsi dan yang telah mengajarkan penulis tata cara menulis karya ilmiah dengan benar,
5. Orangtua dan keluarga yang bersedia menyediakan segala keperluan selama kuliah dan pembuatan laporan skripsi,
6. Albert Irawan, Peter Pangestu, Ricky Jiandy dan Calvina Adrilia selaku sahabat yang selalu memberikan semangat dan segala bantuan dalam awal perkuliahan hingga penyusunan laporan skripsi ini,

7. Pandu Baraja, Julio Christian Young, Eka Jaya Harsono, Alfian Setyowijoyo, Audy, Edo Fernando serta teman-teman lainnya yang selalu memberikan semangat dan segala bantuan dalam penyusunan laporan skripsi ini, dan
8. Seluruh pihak yang tidak mungkin untuk disebutkan satu persatu dalam kata pengantar ini.

Pada akhir kata, penulis menyadari penelitian ini masih jauh dari kata sempurna dan terdapat kekurangan. Oleh karena itu besar harapan untuk berbagai masukan, kritik dan saran yang konstruktif untuk memperbaiki berbagai bentuk kesalahan yang masih ada pada penelitian ini. Dengan terselesaikannya laporan ini, penulis berharap laporan skripsi ini dapat bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi bagi para pembaca.

Tangerang, 22 Agustus 2016

Brinardi Leonardo

UMMN

PERBANDINGAN RABIN-KARP ALGORITHM DAN JARO WINKLER DISTANCE ALGORITHM DALAM PENDETEKSI PLAGIARISME PADA DOKUMEN TEKS

ABSTRAK

Perkembangan teknologi internet semakin pesat yang berakibat semakin banyaknya informasi yang tersedia. Semakin banyak kemudahan yang diberikan semakin memudahkan seseorang dalam melakukan penjiplakan suatu karya. Menurut Colin Neville, *plagiarism* merupakan tindakan atau praktek yang dianggap oleh universitas merupakan suatu kecurangan dengan cara mengambil ide atau tulisan orang lain tanpa menyebutkan rujukan dan diklaim sebagai miliknya. Sistem pendeteksian plagiarisme dalam dokumen teks umumnya mengimplementasikan algoritma pencocokan *string* untuk mencari kesamaan kata (*string*) antar dokumen. Terdapat berbagai macam algoritma yang digunakan untuk pencocokan *string* salah satunya ialah Algoritma Rabin Karp dan Algoritma Jaro Winkler Distance. Algoritma Rabin Karp merupakan salah satu algoritma yang cocok untuk mengatasi permasalahan *multiple string pattern*. Selain itu algoritma ini mampu melakukan komputasi secara efisien dengan memanfaatkan fungsi *hashing*. Dibandingkan dengan algoritma *string matching* lainnya, Algoritma Jaro Winkler Distance memiliki kelebihan dari segi waktu. Algoritma ini memiliki *quadratic runtime complexity* yang sangat efektif dan dapat bekerja pada *string* pendek. Dari permasalahan tersebut, maka dilakukan perbandingan kedua algoritma tersebut. Aplikasi ini dikembangkan berbasis *website* dan digunakan untuk melakukan proses pengujian beberapa tipe-tipe dokumen, meliputi dokumen doc, docx, pdf dan txt. Dari hasil implementasi dan pengujian Algoritma Rabin-Karp dan Algoritma Jaro Winkler Distance didapatkan bahwa kedua algoritma tersebut dapat digunakan untuk melakukan pendeteksian dokumen. Namun dari segi keefektifannya Algoritma Rabin-Karp jauh lebih efektif dan jauh lebih cepat dalam melakukan proses pendeteksian dengan ukuran dokumen yang besar (> 1000 KB).

Kata kunci: Algoritma Rabin-Karp, Algoritma Jaro Winkler Distance, Pencocokan *String*, *Plagiarisme*.

COMPARISON OF RABIN-KARP ALGORITHM AND JARO WINKLER DISTANCE ALGORITHM FOR DETECTION PLAGIARISM IN TEXT DOCUMENTS

ABSTRACT

The development of internet technologies increasingly rapidly which resulted in more information being available. The more conveniences provided even easier for a person in a work of plagiarism. According to Colin Neville, plagiarism is an act or practice that is considered by the university is a fraud by taking someone ideas or writings without mentioning the references and claimed as his own. Plagiarism detection system is generally implemented string matching algorithm in a text document to search for common words (string) between documents. There are a variety of algorithms used for string matching one of them is Algorithm Rabin Karp and Jaro Winkler Distance Algorithm. Rabin Karp algorithm is one of compatible algorithms to solve the problem of multiple string pattern. Besides that, this algorithm is capable of performing computations efficiently by utilizing the functions of hashing. Compared with other string matching algorithm, Jaro Winkler Distance algorithm has advantages in terms of time. This algorithm has quadratic runtime complexity that is very effective and can work on a short string. Of these problems, then do a comparison of the two algorithms. This application are developed based website and the testing process used to make some types of documents, including documents doc, docx, pdf and txt. From the results of the implementation and testing of the Rabin-Karp algorithm and Jaro Winkler Distance algorithm is obtained that both of these algorithms can be used to perform detection of document. But in terms of their effectiveness Rabin Karp algorithm is much more effective and much faster in the process of detecting the document with the size of a large document (> 1000 KB).

Keywords: Rabin-Karp Algorithm, Jaro Winkler Distance Algorithm, String Matching, Plagiarism.

UMN

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN SKRIPSI.....	ii
HALAMAN PERNYATAAN TIDAK PLAGIAT.....	iii
HALAMAN PERSEMBAHAN/MOTO.....	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan Laporan Penelitian	5
BAB II LANDASAN TEORI.....	7
2.1 Plagiarisme.....	7
2.2 Algoritma	8
2.2.1 Kompleksitas Algoritma	9
2.2.2 Performa Algoritma	11
2.3 String Matching	11
2.4 Algoritma Rabin-Karp	13
2.4.1 Konsep Algoritma Rabin-Karp	14
2.5 Algoritma Jaro Winkler Distance	17
2.6 Analisis Perbandingan Rabin-Karp Algorithm dan Jaro Winkler Distance Algorithm.....	19
BAB III METODOLOGI DAN PERANCANGAN SISTEM.....	23
3.1 Metode Penelitian	23
3.2 Perancangan Aplikasi.....	25
3.2.1 Flowchart.....	26
3.2.2 Entity Relationship Diagram (ERD)	34
3.3 Rancangan Tampilan Antarmuka Aplikasi	38
BAB IV IMPLEMENTASI DAN UJI COBA	47
4.1 Spesifikasi Sistem	47
4.2 Implementasi.....	48
4.2.1 Implementasi Algoritma.....	49
4.2.2 Implementasi Program dan Antarmuka	53
4.3 Pengumpulan Sampel Data.....	67
4.4 Analisis Hasil Uji Coba	67
BAB V SIMPULAN DAN SARAN	92
5.1 Simpulan	92
5.2 Saran	93
DAFTAR PUSTAKA	94



UMN

DAFTAR TABEL

Tabel 2.1	Pengelompokan algoritma berdasarkan notasi Big-O	10
Tabel 3.1	Rencana Kegiatan Penelitian.....	23
Tabel 3.2	Struktur Tabel users	35
Tabel 3.3	Struktur Tabel upload.....	36
Tabel 3.4	Struktur Tabel hasil.....	37
Tabel 4.1	Hasil uji coba pada dokumen teks (.txt) dengan ukuran < 1000 KB	68
Tabel 4.2	Hasil uji coba pada dokumen doc atau docx dengan ukuran < 1000 KB	71
Tabel 4.3	Hasil uji coba pada dokumen doc atau docx dengan ukuran < 1000 KB gagal ditemukan kesamaan.....	75
Tabel 4.4	Hasil uji coba pada dokumen pdf dengan ukuran < 1000 KB	76
Tabel 4.5	Hasil uji coba pada dokumen pdf dengan ukuran < 1000 KB gagal ditemukan kesamaan	80
Tabel 4.6	Hasil uji coba pada dokumen yang berukuran > 1000 KB	81
Tabel 4.7	Hasil uji coba pada dokumen yang berukuran > 1000 KB gagal ditemukan kesamaan	84
Tabel 4.8	Hasil uji coba pada dokumen dengan tipe berbeda	86
Tabel 4.9	Hasil uji coba pada dokumen dengan tipe berbeda gagal ditemukan kesamaan.....	89

UMMN

DAFTAR GAMBAR

Gambar 2.1	Algoritma Rabin-Karp (Firdaus, 2008)	14
Gambar 2.2	Pengecekan tiga karakter pertama (Nugroho, 2011)	15
Gambar 2.3	Pengecekan terhadap <i>substring</i> berikutnya (Nugroho, 2011)	15
Gambar 2.4	Pengecekan pattern “c a b” dengan substring “a b b” (Nugroho, 2011)	16
Gambar 2.5	Pembandingan pattern dengan substring (Nugroho, 2011)	16
Gambar 2.6	Pembandingan pattern yang mempunyai nilai hash sama dengan substring (Nugroho, 2011).....	16
Gambar 2.7	Hasil pencarian <i>pattern</i> ditemukan (Nugroho, 2011).....	17
Gambar 3.1	Flowchart Utama Aplikasi.....	26
Gambar 3.2	Flowchart Login	27
Gambar 3.3	Flowchart Register.....	28
Gambar 3.4	Flowchart Member	29
Gambar 3.5	Flowchart Menu Upload Dokumen	30
Gambar 3.6	Flowchart Check Kriteria	30
Gambar 3.7	Flowchart Menu Analisa Dokumen.....	31
Gambar 3.8	Flowchart Submenu Tabel Hasil Analisa	32
Gambar 3.9	Flowchart Algoritma Rabin-Karp.....	32
Gambar 3.10	Flowchart Algoritma Jaro Winkler Distance.....	33
Gambar 3.11	Entity Relationship Diagram	34
Gambar 3.12	Rancangan Utama Interface	40
Gambar 3.13	Rancangan Halaman Login	40
Gambar 3.14	Rancangan Halaman Register.....	41
Gambar 3.15	Rancangan Halaman About	42
Gambar 3.16	Rancangan Halaman Credit	43
Gambar 3.17	Rancangan Halaman Member	43
Gambar 3.18	Rancangan Halaman Upload Dokumen	44
Gambar 3.19	Rancangan Halaman Analisa Dokumen	44
Gambar 3.20	Rancangan Halaman Hasil Analisa	45
Gambar 3.21	Rancangan Halaman About Developer	46
Gambar 4.1	Implementasi Algoritma Rabin Karp	49
Gambar 4.2	Mengambil isi dokumen dan melakukan clean symbol	50
Gambar 4.3	Implementasi Algoritma Jaro Winkler Distance	51
Gambar 4.4	Implementasi Algoritma Jaro Winkler	52
Gambar 4.5	Mencari Nilai Prefiks	53
Gambar 4.6	Tampilan Halaman Home.....	54
Gambar 4.7	Tampilan Halaman Login.....	55
Gambar 4.8	Tampilan Halaman Register	56
Gambar 4.9	Tampilan Halaman Home Setelah Login	56
Gambar 4.10	Tampilan Halaman Upload Dokumen.....	57
Gambar 4.11	Tampilan Halaman Upload Dokumen ketika akan diupload 1 dokumen	58
Gambar 4.12	Tampilan Halaman Upload Dokumen ketika akan diupload lebih dari 1 dokumen	59

Gambar 4.13	Tampilan Halaman Upload Dokumen ketika dokumen yang diupload tidak memenuhi kriteria.....	60
Gambar 4.14	Tampilan Halaman Upload Dokumen ketika dokumen berhasil diunggah	60
Gambar 4.15	Tampilan Halaman Analisa Dokumen	61
Gambar 4.16	Tampilan Halaman File yang Akan Dianalisa.....	62
Gambar 4.17	Tampilan Halaman File yang Akan Dibandingkan	63
Gambar 4.18	Tampilan Halaman Analisa Dokumen ketika tombol Analisis diklik	64
Gambar 4.19	Tampilan Halaman Hasil	64
Gambar 4.20	Tampilan Halaman Show Highlight	65
Gambar 4.21	Tampilan Halaman About	66
Gambar 4.22	Tampilan Halaman Credit	66

