

**METODE *JARO-WINKLER DISTANCE* UNTUK
DETEKSI KEMIRIPAN ABSTRAK SKRIPSI PADA
UNIVERSITAS BHAYANGKARA JAKARTA RAYA**

SKRIPSI

Oleh:

ALANWARI ABBAS SYAHARUDDIN

201910225139

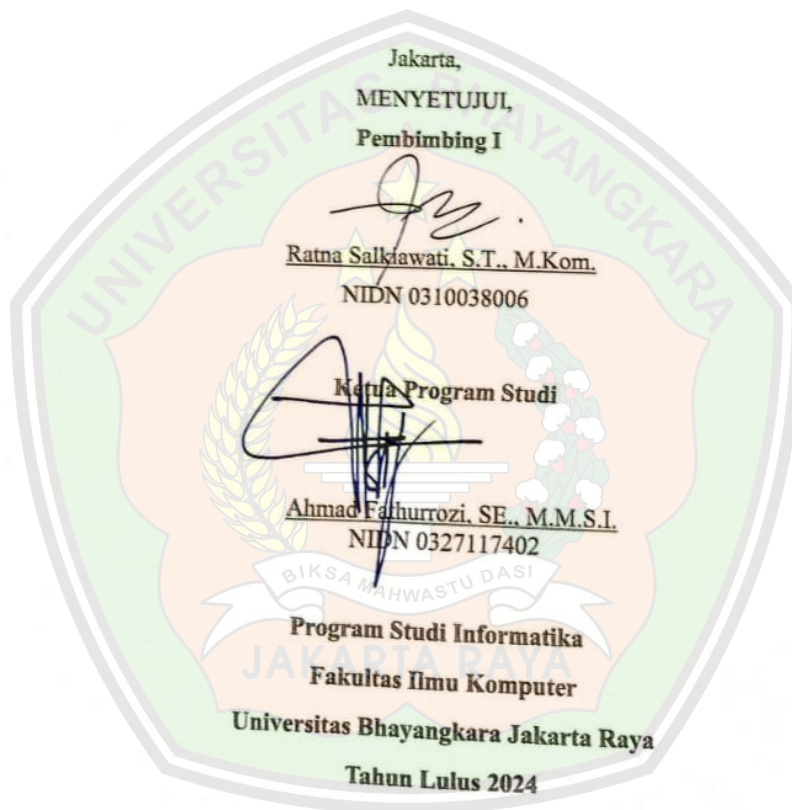


**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA**

2024

LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Tugas akhir : Metode *Jaro-Winkler Distance* Untuk Deteksi Kemiripan Abstrak Skripsi Pada Universitas Bhayangkara Jakarta Raya
Nama Mahasiswa : Al Anwari Abbas Syaharuddin
Nomor Pokok Mahasiswa : 201910225139
Program Studi/Fakultas : Informatika / Ilmu Komputer



LEMBAR PENGESAHAN

Judul Tugas Akhir : Metode *Jaro-Winkler Distance* Untuk Deteksi
Kemiripan Abstrak Skripsi Pada Universitas
Bhayangkara Jakarta Raya
Nama Mahasiswa : Al Anwari Abbas Syaharuddin
Nomor Pokok Mahasiswa : 201910225139
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian : 12 Februari 2024
Tugas Akhir

Jakarta, 12 Februari 2024

MENGESAHKAN,

Ketua Tim Penguji : Mugiarso, S.Kom., M.Kom.
NIDN. 0420117403
Penguji II : Fata Nidaul Khasanah, S.Kom., M.Eng.
NIDN. 0327059202
Penguji III : Ratna Salkiawati, S.T., M.Kom.
NIDN. 0310038006

MENGETAHUI,

Ketua
Program Studi Informatika

Dekan
Fakultas Ilmu Komputer

Ahmad Fahurrozi, S.E., M.M.S.I
NIP. 2012486

Dr. Dra. Tyastuti Sri Lestari, M.M
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA
RAYA FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan di bawah ini :

Nama : Al Anwari Abbas Syaharuddin
NPM : 201910225139
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Metode *Jaro-Winkler Distance* Untuk
Deteksi Kemiripan Abstrak Skripsi Pada
Universitas Bhayangkara Jakarta Raya

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggung jawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 13 Februari
2024 Peneliti



Al Anwari Abbas Syaharuddin

ABSTRAK

AL ANWARI ABBAS SYAHARUDDIN 201910225139, Metode *Jaro-Winkler Distance* untuk deteksi kemiripan abstrak skripsi pada universitas bhayangkara jakarta raya,2024.

Universitas bhayangkara jakarta raya (ubahara jaya) merupakan salah satu perguruan tinggi swasta yang berada dibawah pembinaan yayasan brata bhakti sebagai badan penyelenggara. Skripsi merupakan salah satu persyaratan dasar untuk gelar sarjana. Maka dari itu mahasiswa menganggap skripsi itu sangat sulit. Sehingga banyaknya dokumen abstrak skripsi yang serupa. Berdasarkan masalah yang ada tujuan dari penelitian kali ini adalah menerapkan deteksi kemiripan abstrak skripsi menggunakan metode jaro-winkler distance untuk mengetahui nilai dari similarity. Langkah yang digunakan pada metode ini yaitu menghitung karakter yang sama, lalu mencari nilai transposisi, kemudian menghitung nilai dari jaro winkler distance dan hasil nilai similarity. Hasil dari penelitian ini adalah menentukan nilai similarity nya yaitu: jika kedua dokumen di bandingkan dan mendapatkan hasil sebesar 83%, maka kedua dokumen yang di bandingkan mirip, sedangkan bila kedua dokumen di bandingkan dan mendapatkan hasil 0% berarti kedua dokumen tersebut tidak terdapat kemiripan sama sekali.

Kata kunci: Algoritma *Jaro-Winkler Distance*, Deteksi Kemiripan, Abstrak Skripsi.

ABSTRACT

AL ANWARI ABBAS SYAHARUDDIN 201910225139, *Jaro-Winkler Distance Method for detecting abstract similarity of the thesis at Bhayangkara University Jakarta Raya, 2024. Bhayangkara University Jakarta Raya (ubahara jaya) is one of the private higher education institutions under the auspices of the Brata Bhakti foundation as the organizing body. A thesis is one of the basic requirements for a bachelor's degree. Therefore, students consider the thesis to be very difficult. As a result, there are many similar abstract documents of the thesis. Based on the existing problems, the aim of this research is to apply the detection of abstract thesis similarity using the Jaro-Winkler distance method to determine the value of similarity. The steps used in this method are counting the same characters, finding the transposition value, then calculating the value of the Jaro-Winkler distance and the resulting similarity value. The result of this research is to determine its similarity value: if the two documents are compared and achieve a result of 83%, then the two documents are considered similar, while if the two documents are compared and get a result of 0%, it means that the two documents have no similarity at all.*

Keywords: *Jaro-Winkler Distance Algorithm, Similarity Detection, Thesis Abstract.*

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA
ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas
Bhayangkara Jakarta Raya, sayayang bertanda
tangan di bawah ini :

Nama : Al Anwari Abbas Syaharuddin
NPM : 201910225139
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non- Eksklusif (Non- Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul:

**METODE JARO-WINKLER DISTANCE UNTUK DETEKSI
KEMIRIPAN ABSTRAK SKRIPSI PADA UNIVERSITAS
BHAYANGKARA JAKARTA RAYA**

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non- eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalih mediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 13 Februari 2024
Yang Meny

Al Anwari Abbas Syaharuddin



KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga saya dapat menyusun skripsi yang berjudul “METODE *JARO-WINKLER DISTANCE* UNTUK DETEKSI KEMIRIPAN ABSTRAK SKRIPSI PADA UNIVERSITAS BHAYANGKARA JAKARTA RAYA”.

Dalam penyusunan Skripsi ini, penulis dengan segala hormat serta kerendahan hati, perkenankanlah penulis mengucapkan terima kasih kepada :

1. Allah SWT atas limpahan Rahmat dan berkah-Nya bagi penulis hingga saat ini. Selain itu, Nabi Muhammad SAW adalah pemimpin umat manusia dari era jahiliyah ke era Islamiyah.
2. Kepada kedua Orang Tua yang telah membesarkan dan mendidik penulis memberikan semangat dan motivasi terbesar dalam kehidupan.
3. Bapak Irjen Pol. (Purn) Drs. H. Bambang Karsono, S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
4. Dr. Dra. Tyastuti Sri Lestari, MM. Selaku Dekan Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
5. Bapak Ahmad Fathurozi, S.E.,M.M.S.I selaku Kepala Prodi Informatika Universitas Bhayangkara Jakarta Raya.
6. Ibu Ratna Salkiawati, S.T., M.Kom selaku Dosen Pembimbing Skripsi yang sangat membantu dalam memberi masukan serta motivasi untuk penulis selama penelitian berlangsung.

7. Ibu Dwi Budi Srisulistiowati, S.Kom., M.M selaku dosen pembimbing Akademik.
8. Muhammad Syahrullah, Bangun Tegar, Jasmine Kinari, Inesyah Irena, dan seluruh teman teman saya yang tidak bisa sebutkan satu persatu yang telah membantu dan memberikan saya masukan dan dukungan.
9. Puti Wulandari Dan Nahdillah selaku Kakak dan Adik saya yang telah membantu dan memberikan *Support* serta dukungan motivasi untuk menyelesaikan penelitian ini. Semoga Allah SWT selalu memeberikan perlindungan-Nya.

Penulis menyadari bahwa masih banyak kekurangan dari proposal kripsi ini, baik dari materi maupun dari segi teknik penyajiannya. Oleh karena itu, penulis mengharapkan kritik serta saran sangat penulis terima dengan senang hati. Akhir kata Penulisan berharap agar skripsi ini dapat memberikan manfaat yang sangat besar.

Jakarta,

2024



Al Anwari Abbas Syaharuddin

201910225139

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iv
ABSTRAK	v
<i>ABSTRAK</i>.....	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASIKARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan masalah.....	4
1.4 Batasan Masalah.....	4
1.5 Tujuan dan Manfaat Penelitian.....	5
1.5.1 Tujuan Penelitian.....	5

1.5.2	Manfaat Penelitian	5
1.6	Sistematika Penulisan.....	5
BAB II LANDASAN TEORI		7
2.1	Tinjauan Pustaka.....	7
2.2	<i>Natural Language Processing(NLP)</i>	8
2.3	<i>Text Preprocessing</i>	9
2.3.1	<i>Case Folding</i>	9
2.3.2	Tokenizing.....	9
2.3.3	<i>Stopword</i>	10
2.3.4	<i>Stemming</i>	10
2.4	<i>Jaro-Winkler Distance</i>	11
2.5	<i>Similarity</i>	12
2.6	Alat Bantu Perangkat Lunak Pendukung Pemrograman	14
2.6.1	Pengertian <i>flask</i>	14
2.6.2	<i>Python</i>	14
2.6.3	HTML (<i>Hypertext Markup Language</i>)	14
2.6.4	HeidiSQL	15
2.6.5	Laragon	15
BAB III METODOLOGI PENELITIAN		16
3.1	Tempat dan Waktu Penelitian	16

3.2	Struktur Organisasi	17
3.3	Kerangka Penelitian	20
3.4	Perangkat Keras (<i>Hardware</i>)	22
3.5	Perangkat Lunak (<i>Software</i>)	22
3.6	Analisa Sistem Berjalan	22
3.7	Sistem Usulan.....	24
BAB IV HASIL DAN PEMBAHASAN.....		28
4.1	Pembahasan	28
4.2	Input.....	28
4.2.1	<i>Text Preprocessing</i>	31
4.2.1.1	Case Folding	32
4.2.1.2	Tokenizing	32
4.2.1.3	Stopword removal	33
4.2.1.4	Stemming.....	33
4.3	<i>Jaro-Winkler Distance</i>	34
4.4	Implementasi	36
4.5	Pengujian	39
4.6	Hasil Uji Coba	40
BAB V PENUTUP		42
5.1	Kesimpulan.....	42

5.2	Saran.....	42
DAFTAR PUSTAKA.....		43



DAFTAR TABEL

Tabel 1.1 contoh dataset abstrak skripsi.....	3
Tabel 2. 1 Tinjauan Pustaka	7
Tabel 2. 2 <i>Case Folding</i>	9
Tabel 2. 3 <i>Tokenizing</i>	10
Tabel 2. 4 <i>Stopword</i>	10
Tabel 2. 5 <i>Stemming</i>	11
Tabel 3. 1 Waktu dan Kegiatan Penelitian	19
Tabel 3. 2 Range persentase similarity.....	26
Tabel 4. 1 Dataset Abstrak Skripsi	29
Tabel 4. 2 Abstrak Skripsi	31
Tabel 4. 3 <i>case folding</i>	32
Tabel 4. 4 <i>Tokenizing</i>	32
Tabel 4. 5 <i>stopword removal</i>	33
Tabel 4. 6 <i>stemming</i>	33
Tabel 4. 7 <i>Input</i>	35
Tabel 4. 8 jumlah masing-masing dokumen	35
Tabel 4. 9 jumlah <i>character</i>	35
Tabel 4. 10 Pengujian Metode <i>Jaro winkler</i>	39
Tabel 4. 11 Hasil Perhitungan <i>Jaro-winkler Distance</i>	41

DAFTAR GAMBAR

Gambar 1. 1 Skripsi Informatika.....	3
Gambar 2. 1 <i>Text Preprocessing</i>	9
Gambar 3. 1 Alamat	16
Gambar 3. 2 Struktur Organisasi Perpustakaan Bhayangkara jakarta Raya	17
Gambar 3. 3 Kerangka Penelitian	20
Gambar 3. 4 Flowchart Sitem berjalan	23
Gambar 3. 5 Sistem Usulan.....	24
Gambar 4. 1 Tampilan Dataset didalam Database	30
Gambar 4. 2 Hasil dari Tabel kata.....	30
Gambar 4. 3 <i>Text Preprocessing</i>	31
Gambar 4. 4 <i>Flowchart Jaro Winkler Distance</i>	34
Gambar 4. 5 menjalankan program.....	37
Gambar 4. 6 tampilan untuk menginput query.....	37
Gambar 4. 7 Hasil <i>Score jaro-winkler distance</i>	38
Gambar 4. 8 Hasil bila tidak ada kesamaan	39

DAFTAR LAMPIRAN

Lampiran 1 Kartu bimbingan Skripsi.....	47
Lampiran 2 Kartu bimbingan.....	48
Lampiran 3 Surat Izin Penelitian	48
Lampiran 4 Hasil Plagiarism Checker	50
Lampiran 5 Biodata.....	51
Lampiran 6 Dataset.....	52

