



**IMPLEMENTASI ALGORITMA COSINE SIMILARITY DENGAN  
METODE TF-IDF UNTUK PENGECEKAN JUDUL SKRIPSI  
PROGRAM STUDI INFORMATIKA UNIVERSITAS MERCU  
BUANA**

LAPORAN SKRIPSI

MUHAMMAD RASYID  
41518310002

UNIVERSITAS  
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA JAKARTA  
2022**



**IMPLEMENTASI ALGORITMA COSINE SIMILARITY DENGAN  
METODE TF-IDF UNTUK PENGECEKAN JUDUL SKRIPSI  
PROGRAM STUDI INFORMATIKA UNIVERSITAS MERCU  
BUANA**

LAPORAN SKRIPSI

MUHAMMAD RASYID  
41518310002

UNIVERSITAS  
MERCU BUANA

Diajukan sebagai salah satu syarat untuk  
memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA JAKARTA  
2022**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Rasyid  
NIM : 41518310002  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Cosine Similarity Dengan Metode TF - IDF Untuk Pengecekan Judul Skripsi Program Studi Informatika Universitas Mercu Buana

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS  
MERCU BUANA

Jakarta, 27 Desember 2022



Muhammad Rasyid

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Rasyid  
NIM : 41518310002  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Implementasi Algoritma Cosine Similarity Dengan Metode TF-IDF Untuk Pengecekan Judul Skripsi Program Studi Informatika Universitas Mercu Buana

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Leonard Goeirmanto, Dr., ST, M.Sc  
NIDN : 0312087601  
Ketua Penguji : Eliyani, Dr. Ir.  
NIDN : 0321026901  
Penguji 1 : Rahmat Rian Hidayat, ST, MMSI  
NIDN : 0317049002  
Penguji 2 : Dwi Anindyani Rocmah, ST, MTI  
NIDN : 0011057801


( 28.03.2023 )  
  
  


Jakarta, 6 Maret 2023

Mengetahui,

Koordinator Tugas Akhir

Ketua Program Studi

  
Wawan Gunawan, S.Kom, MT

  
Ir. Emil R. Kaburuan, Ph.D., IPM.

## KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana
2. Bapak Yaya Sudarya Triana, M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer
3. Bapak Ir. Emil R. Kaburuan, Ph.D., IPM. selaku Ketua Program Studi Teknik Informatika
4. Bapak Leonard Goeirmanto, Dr., ST, M.Sc selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini
5. Kedua Orang Tua Peneliti yang selalu mendukung penulis, serta mendoakan kelancaran terbuatnya laporan ini.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 27 Desember 2022

Penulis

## HALAMAN PERNYATAAN PERSUTUJUAN PUBLIKASI KARYA ILMIAH

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Rasyid  
NIM : 41518310002  
Program Studi : Teknik Informatika  
Judul Tugas Akhir : Implementasi Algoritma Cosine Similarity Dengan Metode TF - IDF Untuk Pengecekan Judul Skripsi Program Studi Informatika Universitas Mercu Buana

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 27 Desember 2022

Yang menyatakan,

  
METERAI TEMPEL  
13194AKX309154302  
Muhammad Rasyid

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PERNYATAAN KARYA SENDIRI</b> .....	ii
<b>HALAMAN PENGESAHAN</b> .....	iii
<b>KATA PENGANTAR</b> .....	iv
<b>HALAMAN PERNYATAAN PERSUTUJUAN PUBLIKASI KARYA ILMIAH</b> ..	v
<b>ABSTRAK</b> .....	vi
<b>ABSTRACT</b> .....	vii
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR TABEL</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>BAB I PENDAHULUAN</b> .....	1
1.1. Latar Belakang .....	1
1.2. Perumusan Masalah .....	3
1.3. Tujuan Penelitian .....	3
1.4. Manfaat Penelitian .....	3
1.5. Batasan Penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA</b> .....	5
2.1. Algoritma Pemrograman.....	5
2.2. Rekayasa Perangkat Lunak.....	14
2.3. Sistem Basis Data .....	17
<b>BAB III METODE PENELITIAN</b> .....	23
3.1. Jenis Penelitian .....	23
3.2. Tahapan Penelitian.....	23
<b>BAB IV PEMBAHASAN</b> .....	26
4.1. Definisi Kebutuhan .....	26
4.2. Analisis Kebutuhan.....	28
4.3. Disain Sistem .....	29
4.4. <i>User Interface</i> .....	32
4.5. Implementasi Sistem.....	33
4.6. Pengujian Sistem.....	47
<b>BAB V KESIMPULAN DAN SARAN</b> .....	49
5.1. Kesimpulan .....	49
5.2. Saran .....	49
<b>DAFTAR PUSTAKA</b> .....	50
<b>LAMPIRAN</b> .....	53

## DAFTAR TABEL

Tabel 4. 1 Hasil drop data duplicate .....	35
Tabel 4. 2 Output Stopword .....	38
Tabel 4. 3 Output nilai TF dataset .....	39
Tabel 4. 4 Output nilai IDF dataset .....	40
Tabel 4. 5 Output nilai TF query .....	41
Tabel 4. 6 Output nilai IDF query .....	42
Tabel 4. 7 Nilai TF-IDF query .....	43
Tabel 4. 8 Rank Similarity.....	46





## DAFTAR GAMBAR

Gambar 2. 1 Flowchart algoritma bubble sort.....	5
Gambar 2. 2 Script bubble sort.....	8
Gambar 2. 3 Tampilan hasil algoritma bubble sort .....	8
Gambar 2. 4 Flowchart algoritma selection sort.....	9
Gambar 2. 5 Script selection sort .....	13
Gambar 2. 6 Tampilan hasil algoritma selection sort.....	14
Gambar 2. 7 Tahapan metodologi waterfall .....	14
Gambar 2. 8 Tahapan metodologi agile .....	16
Gambar 2. 9 Normal form (Unnormalize).....	20
Gambar 2. 10 Normalisasi 1NF .....	20
Gambar 2. 11 Normalisasi 2NF .....	21
Gambar 2. 12 Normalisasi 3NF.....	21
Gambar 3. 1 Tahapan Penelitian.....	23
Gambar 4. 1 Pembuatan model.....	29
Gambar 4. 2 UI.....	32
Gambar 4. 3 Import library.....	33
Gambar 4. 4 Import dataset .....	34
Gambar 4. 5 Drop data yang tidak dibutuhkan.....	34
Gambar 4. 6 Cek dan hapus data duplikat.....	35
Gambar 4. 7 Input query.....	36
Gambar 4. 8 Tokenizing .....	36
Gambar 4. 9 Tampilan output tokenizing.....	37
Gambar 4. 10 Stopword menggunakan Sastrawi.....	37
Gambar 4. 11 Menentukan nilai TF dari dataset .....	38
Gambar 4. 12 Menentukan nilai IDF dari dataset .....	39
Gambar 4. 13 Menentukan nilai TF dari query .....	41
Gambar 4. 14 Menentukan nilai IDF dari query .....	42
Gambar 4. 15 Menentukan nilai TF-IDF dari query .....	43
Gambar 4. 16 Menentukan nilai cosine similarity.....	44
Gambar 4. 17 Ranked hasil perhitungan cosine similarity.....	45
Gambar 4. 18 Tampilan menjalankan file python .....	47
Gambar 4. 19 Tampilan memasukkan input query judul .....	47
Gambar 4. 20 Tampilan hasil perhitungan cosine similarity.....	48

## DAFTAR LAMPIRAN

Lampiran 1 Surat Pernyataan .....	53
Lampiran 2 Surat Pengalihan Hak Cipta .....	54
Lampiran 3 Scan KTP .....	55
Lampiran 4 Curriculum Vitae.....	56
Lampiran 5 Hasil Turnitin .....	57
Lampiran 6 Form Wawancara 1 .....	58
Lampiran 7 Form Wawancara 2 .....	59
Lampiran 8 Naskah Artikel Jurnal .....	60

