

**PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL
TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN
ALGORITMA COSINE SIMILARITY**

TUGAS AKHIR



diajukan oleh:

Nama : Fadhilah Hafidz Pradana

NIM : 22.01.4769

**PROGRAM DIPLOMA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL
TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN
ALGORITMA COSINE SIMILARITY**

TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat mencapai gelar Ahli Madya
Komputer Program Diploma – Program Studi Teknik Informatika



diajukan oleh

Nama : Fadhilah Hafidz Pradana

NIM : 22.01.4769

Kepada

**PROGRAM DIPLOMA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

TUGAS AKHIR

PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN ALGORITMA COSINE SIMILARITY

yang dipersiapkan dan disusun oleh

Fadhilah Hafidz Pradana

22.01.4769

Telah disetujui oleh Dosen Pembimbing Tugas Akhir
pada tanggal 25 Juni 2025

Dosen Pembimbing,



Barka Satya, M.Kom

NIK. 190302126

HALAMAN PENGESAHAN

TUGAS AKHIR

**PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL
TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN
ALGORITMA COSINE SIMILARITY**

yang disusun dan diajukan oleh

Fadhilah Hafidz Pradana

22.01.4769

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Juni 2025

Susunan Dewan Penguji

Nama Penguji

Hastari Utama, S.Kom., M.Cs.
NIK. 190302230

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Tanda Tangan



Tugas Akhir ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Ahli Madya komputer
Tanggal 19 Juni 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fadhilah Hafidz Pradana
NIM : 22.01.4769

Menyatakan bahwa Tugas Akhir dengan judul berikut:

PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN ALGORITMA COSINE SIMILARITY

Dosen Pembimbing : Barka Satya, M.Kom

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, 19 Juni 2025

Yang Menyatakan,



Fadhilah Hafidz Pradana

HALAMAN PERSEMBAHAN

Bismillahirrahmanirrahim. Dengan segala puji syukur ke hadirat Allah SWT atas segala limpahan rahmat serta hidayahnya sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik. sebagai penulis ingin mengucapkan rasa terima kasih kepada orang-orang yang telah membantu penulis dalam menyelesaikan tugas akhir ini:

1. Kedua orang tua tercinta, ayah ibu terima kasih telah senantiasa mendukung, mendoakan, memberi semangat, serta segala pengorbanan yang telah diberikan sejak awal hingga saat ini. Tugas akhir ini adalah hasil dari perjuangan dan doa beliau yang tak pernah putus.
2. Untuk adik, terima kasih telah mendukung kakak dalam penyusunan tugas akhir ini.
3. Teman-teman serta sahabat, terima kasih atas kebersamaan, dukungan, serta semangat yang kalian berikan.

KATA PENGANTAR

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang dan segala puji syukur atas kehadiran Allah SWT karena telah melimpahkan berkah dan rahmat-Nya. Sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “PENGEMBANGAN SISTEM DETEKSI KEMIRIPAN JUDUL TUGAS AKHIR DAN SKRIPSI MENGGUNAKAN ALGORITMA COSINE SIMILARITY” dengan baik.

Penyusunan tugas akhir ini untuk memenuhi syarat dalam mencapai gelar Ahli Madya komputer dari program studi Teknik Informatika, Fakultas Ilmu Komputer di Universitas Amikom Yogyakarta.

Dalam penyusunan tugas akhir ini penulis mendapat banyak bimbingan dan bantuan, maka pada kesempatan ini penulis ingin mengucapkan terima kasih sebesar-besarnya kepada:

1. Kedua orang tua yang selalu memberikan dukungan dan doa dalam kelancaran penyusunan tugas akhir ini.
2. Prof. Dr. M.Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
3. Prof. Dr. Kusri, M.Kom. Selaku Dekan Fakultas Ilmu Komputer.
4. Bapak Barka Satya, M.Kom. Selaku dosen pembimbing dan Ketua Program Studi D3 Teknik Informatika yang telah memberikan waktu, ilmu, dan bimbingan selama proses penyusunan tugas akhir ini.
5. Seluruh dosen Fakultas Ilmu Komputer yang telah memberikan ilmu selama masa perkuliahan.

Semoga segala kebaikan yang diberikan mendapatkan balasan dari Allah SWT. Penulis berharap sistem deteksi kemiripan judul tugas akhir dan skripsi ini dapat membantu banyak mahasiswa yang ingin mengajukan judul. Jika ada kesalahan penulis mengharapkan kritik dan saran yang membangun untuk memperbaiki kedepannya.

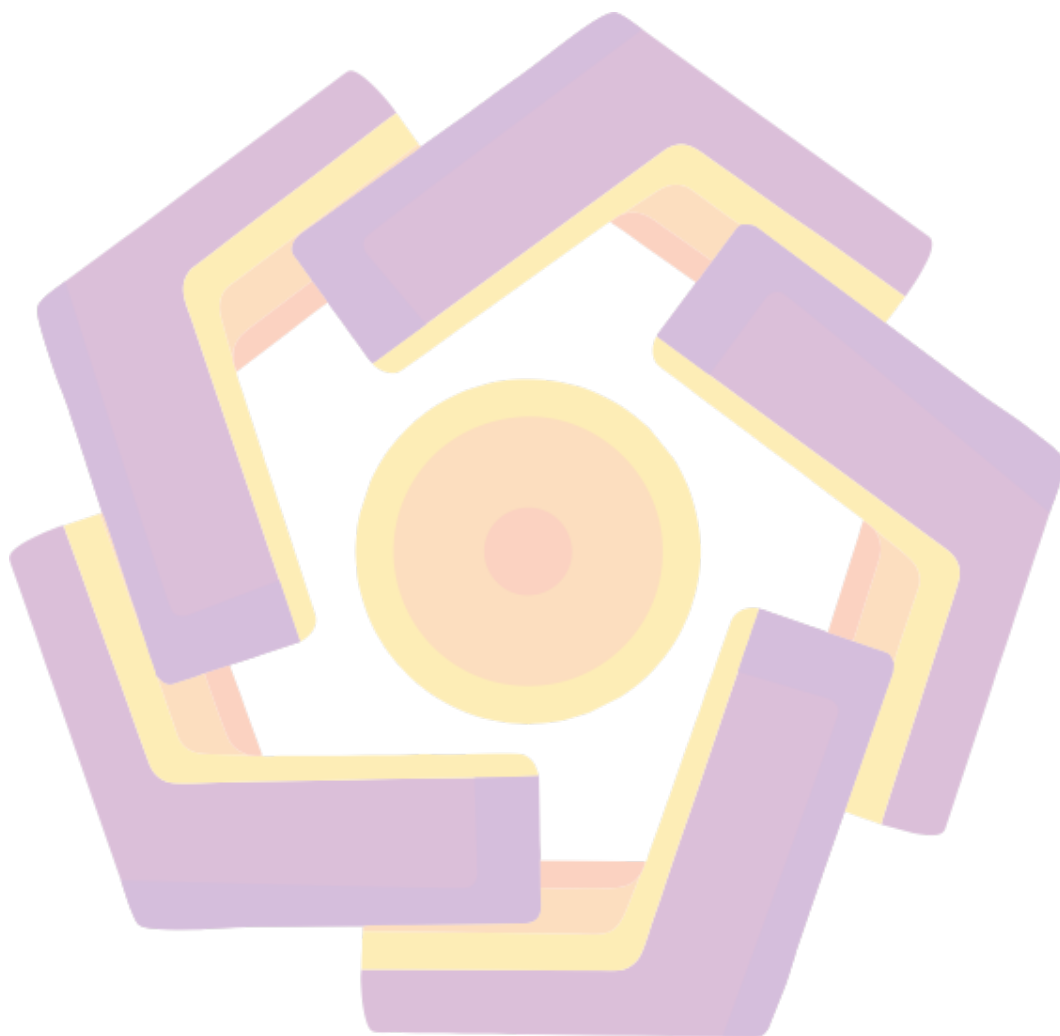
Yogyakarta, 25 Mei 2025

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN.....	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	ix
DAFTAR LAMPIRAN.....	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH	xii
INTISARI	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan masalah.....	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah	2
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 <i>Literature Review</i>	4
2.2 Landasan Teori.....	8
BAB III METODOLOGI PENELITIAN	15
3.1 Pengumpulan Kebutuhan	15
3.2 Langkah Penelitian.....	17
BAB IV HASIL DAN PEMBAHASAN	22
4.1 Implementasi.....	22
4.2 Pengujian.....	26
BAB V KESIMPULAN DAN SARAN	29
5.1 Kesimpulan	29
5.2 Saran	29

DAFTAR PUSTAKA	30
LAMPIRAN.....	33

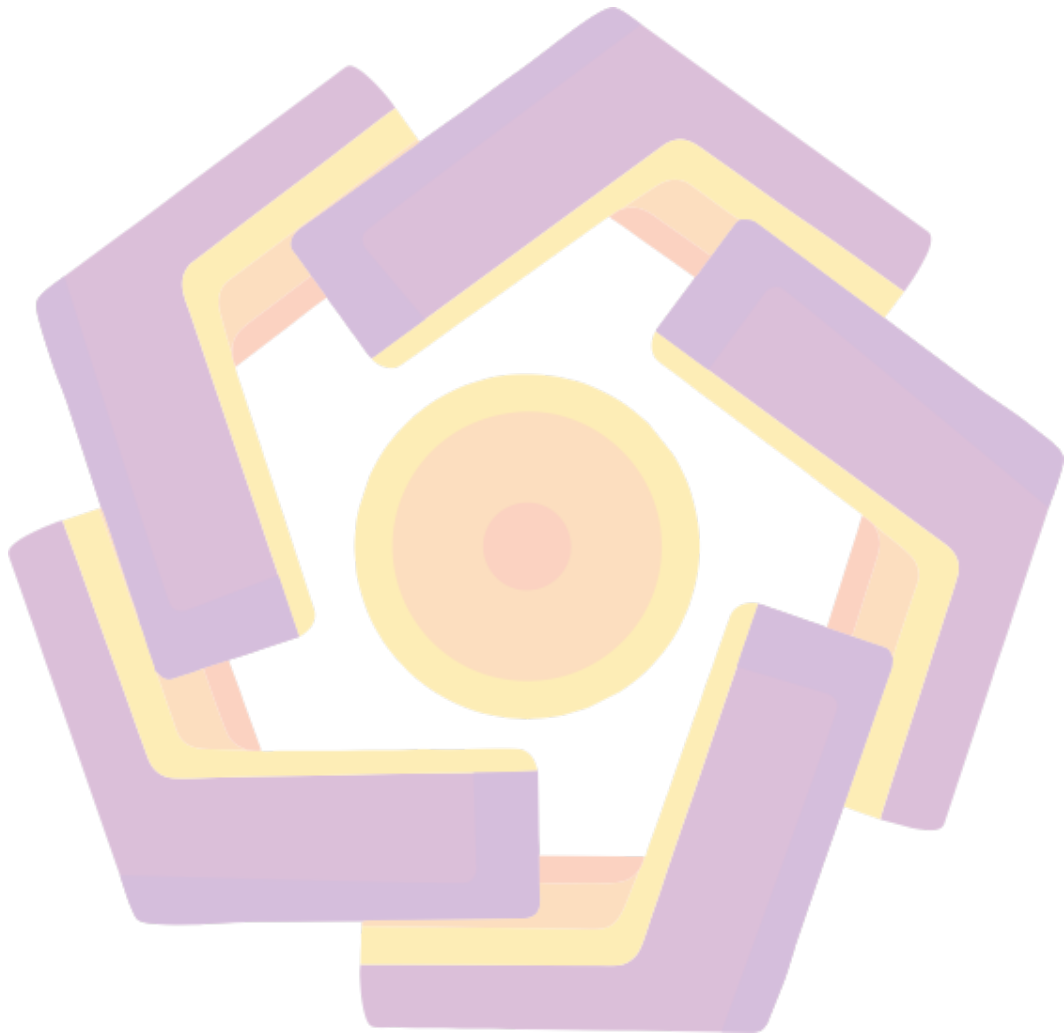


DAFTAR GAMBAR

Gambar 2. 1 Bagan Flowchart	12
Gambar 2. 2 Simbol Use Case Diagram	13
Gambar 3.1 Flowchart sistem deteksi kemiripan judul.....	15
Gambar 3.2 Alur data preprocessing.....	16
Gambar 3.3 Alur pembobotan TF-IDF	17
Gambar 3.4 Tahap metode waterfall.....	18
Gambar 3.5 Use case diagram pengguna dan admin	19
Gambar 3.6 ERD sistem deteksi kemiripan judul.....	20
Gambar 4.1 Halaman awal.....	22
Gambar 4.2 Halaman cek kemiripan.....	22
Gambar 4.3 Halaman hasil cek kemiripan	23
Gambar 4. 4 Perbedaan Warna Skor Kemiripan.....	23
Gambar 4.5 Halaman masuk admin.....	24
Gambar 4.6 Halaman daftar admin.....	24
Gambar 4.7 Halaman dashboard admin.....	25
Gambar 4.8 Halaman data judul	25
Gambar 4. 9 Isi Basisdata	26

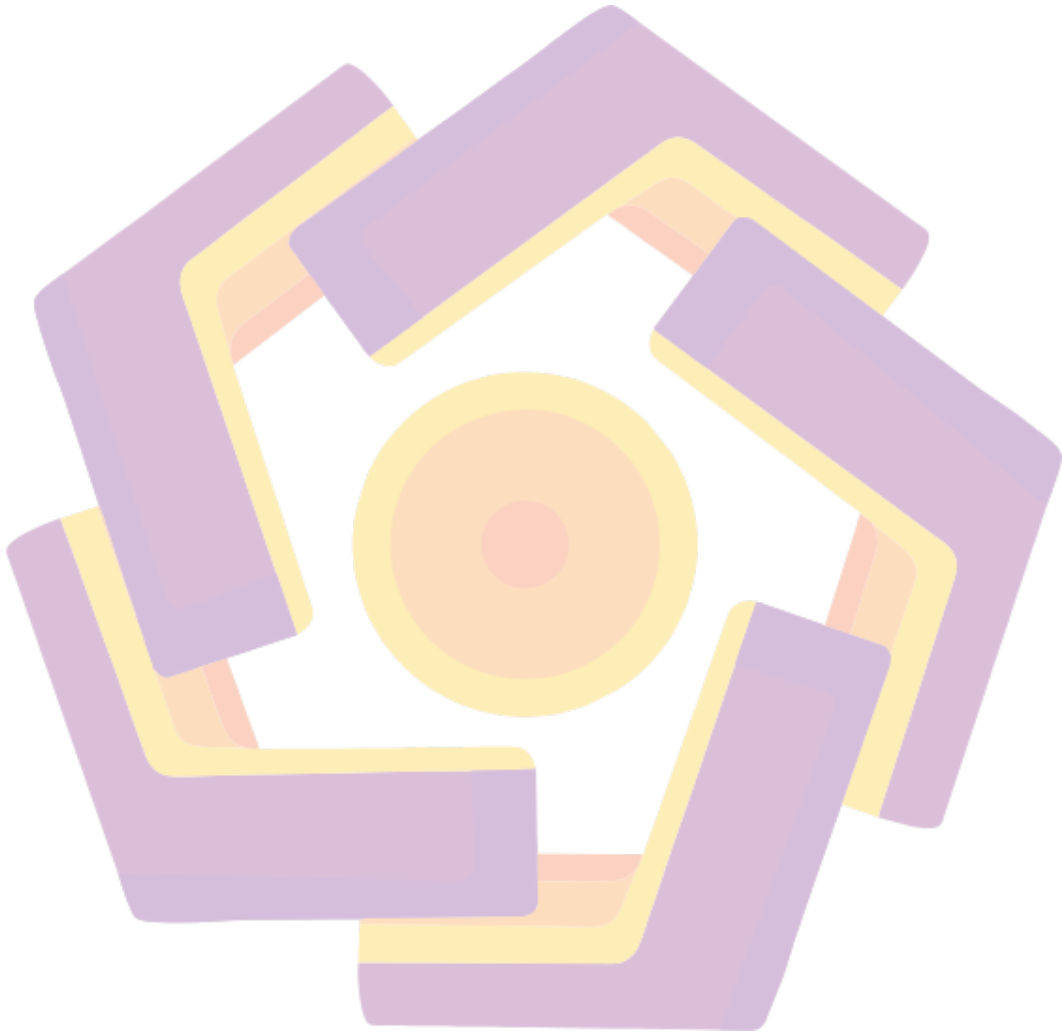
DAFTAR TABEL

Tabel 2.1 Tabel Tinjauan Literatur	6
Tabel 3.1 Tabel Kebutuhan Sistem	18
Tabel 4. 1 Pengujian Blackbox	26



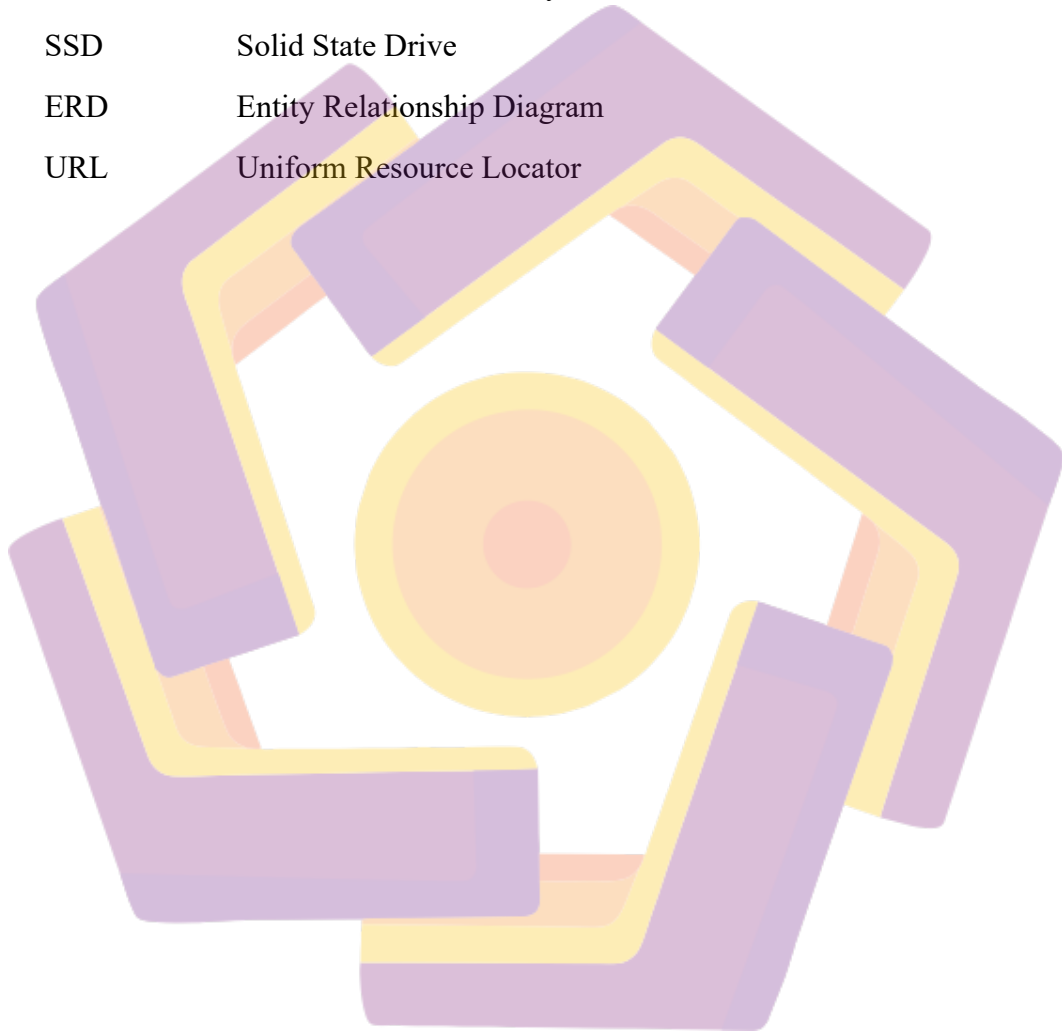
DAFTAR LAMPIRAN

Lampiran 1.1 Link Website.....	33
--------------------------------	----



DAFTAR LAMBANG DAN SINGKATAN

TF-IDF	Term Frequency-Inverse Document Frequency
WWW	World Wide Web
XML	Extensible Markup Language
SQL	Structured Query Language
RAM	Random Access Memory
SSD	Solid State Drive
ERD	Entity Relationship Diagram
URL	Uniform Resource Locator



DAFTAR ISTILAH

<i>Vektor</i>	besaran yang mempunyai arah untuk dibandingkan
<i>Plagiarisme</i>	tindakan meniru karya milik orang lain
<i>Scraping</i>	proses mengambil data dari suatu internet atau website
<i>Term</i>	kata dalam TF-IDF
<i>Query</i>	permintaan informasi atau instruksi ke basisdata
<i>Repository</i>	tempat penyimpanan data
<i>Prototype</i>	rancangan awal sebuah sistem
<i>Copy paste</i>	menyalin dan menempelkan data ke lokasi lain
<i>Relevan</i>	berkaitan dengan suatu hal
<i>Outcome</i>	hasil dari suatu produk yang dikembangkan
<i>Server</i>	sistem yang dirancang khusus untuk menyediakan layanan
<i>Flowchart</i>	alur atau bagan proses dari sebuah sistem
<i>Array</i>	struktur data yang menyimpan kumpulan data
<i>Programming</i>	proses merancang, menulis, menguji kode program
<i>Framework</i>	kerangka atau struktur kerja pada bahasa pemrograman
<i>Developer</i>	orang yang mengembangkan sebuah sistem
<i>Package</i>	cara untuk mengelompokkan dan mengorganisir modul
<i>Library</i>	pustaka yang ada dalam bahasa pemrograman
<i>Middleware</i>	perantara yang berfungsi sebagai penghubung sistem
<i>Cookie</i>	menyimpan informasi kunjungan website
<i>Session</i>	periode waktu ketika pengguna berinteraksi dengan sistem
<i>Security header</i>	keamanan dalam proses integrasi antar website
<i>Data preprocessing</i>	pemrosesan suatu data dalam analisis model
<i>Stemming</i>	proses menghilangkan imbuhan kata atau sinonim
<i>Filter</i>	menyaring data yang tidak diinginkan dari suatu proses

INTISARI

Pada dunia pendidikan era digital yang berkembang sangat pesat ini, terdapat banyak platform-platform digital yang menyediakan referensi pembelajaran bagi mahasiswa. Seperti website google scholar yang menyediakan berbagai jurnal, buku, tesis, artikel, tugas akhir, dan skripsi. Dengan adanya platform digital tersebut dapat menjadi acuan oleh mahasiswa dalam memilih judul tugas akhir dan skripsi yang relevan dan menghindari *plagiarisme*. Namun masih terdapat banyak mahasiswa memiliki judul yang hampir sama dengan mahasiswa lainnya. Oleh karena itu diperlukan adanya sebuah sistem yang dapat memberikan informasi mengenai tingkat kemiripan judul dan dapat menjadi referensi bagi mahasiswa dalam menentukan judul yang unik, relevan, dan tidak serupa dengan mahasiswa lainnya.

Sistem ini dikembangkan menggunakan bahasa pemrograman JavaScript, framework Express untuk backend dan framework ReactJS untuk frontend, serta basisdata MySQL. Sistem ini dikembangkan menggunakan metode Waterfall, di mana data judul diambil dari situs <https://eprints.amikom.ac.id> yang merupakan repository Universitas Amikom Yogyakarta dengan menggunakan teknik *Web Scraping*. Untuk menentukan tingkat kemiripan judul, sistem menerapkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai pembobotan kata dan algoritma *Cosine Similarity* sebagai alat ukur kemiripan judul.

Hasil akhir dalam pengembangan sistem deteksi kemiripan judul ini adalah sebuah aplikasi berbasis website yang memiliki fitur untuk mendeteksi kemiripan judul tugas akhir dan skripsi yang dapat membantu mahasiswa dalam menentukan judul tugas akhir atau skripsi yang relevan dan mengurangi kemiripan dengan judul lainnya.

Kata kunci: Tugas Akhir, Skripsi, *Web Scraping*, TF-IDF, Algoritma *Cosine Similarity*

ABSTRACT

In this rapidly evolving digital education era, there are many digital platforms that provide learning references for students. One such platform is Google Scholar, which provides various journals, books, articles, final assignments, and theses. These digital platforms can be a reference for students in choosing relevant final project and thesis titles and avoiding plagiarism. However, there are still many students who have titles that are almost the same as other students. Therefore, it is necessary to have a system that can provide information about the level of similarity of titles and can be a reference for students in determining titles that are unique, relevant, and not similar to other students.

This system was developed using JavaScript programming language, Express framework for backend and ReactJS framework for frontend, and MySQL database. This system was developed using the Waterfall method, where the title data was taken from the site <https://eprints.amikom.ac.id> which is the repository of Amikom University Yogyakarta using the Web Scraping technique. To determine the level of title similarity, the system applies the Term Frequency-Inverse Document Frequency (TF-IDF) method as a word weighting and the Cosine Similarity algorithm to measure title similarity.

The final result in the development of this title similarity detection system is a web-based application that has features to detect similarities in the title of the final project and thesis that can help students in determining the title of the final project or thesis that is relevant and reduce similarities with other titles.

Keyword: Final Project, Thesis, Web Scraping, TF-IDF, Cosine Similarity Algorithm