



Aplikasi Smart Title Checker Laporan Mahasiswa Menggunakan Metode Cosine Similarity Berbasis Website Pada Politeknik Xyz

Yudistira Sira Permana^{1*}, Muhammad Syahril², Agus Setiawan³

¹Universitas Siguntang Mahaputra, Palembang, Indonesia, Indonesia

²Universitas Muhammadiyah Palembang, Palembang, Indonesia, Indonesia

³Manajemen Informatika, Polikitenik Negeri Sriwijaya, Palembang, Indonesia

Informasi Artikel

Sejarah Artikel:
Submit: 13 May 2026
Revisi: 04 Juni 2026
Diterima: 14 Juni 2026
Diterbitkan: 10 Juli 2026

Kata Kunci

Smart Title Checker, Cosine Similarity, Kemiripan Judul, Validasi Akademik, Website.

Korespondensi

E-mail: yudistirasirapermana@gmail.com

A B S T R A C T

The increasing number of student final reports each year presents challenges for higher education institutions in maintaining the uniqueness and novelty of proposed titles. The title-checking process, which is still commonly conducted manually, tends to require considerable time and may lead to inconsistencies in the validation process. This study aims to design and develop a web-based Smart Title Checker application capable of detecting the similarity level of student report titles using the Cosine Similarity method. This method is applied to measure the similarity between titles, thereby supporting a more objective validation process. The system development was carried out through several stages, including requirements analysis, design, implementation, and application testing. The results show that the application can perform title similarity checking quickly and accurately by displaying similarity percentages against stored title data in the database. The system implementation also helps students obtain initial information regarding potential title similarities and improves the efficiency of the validation process conducted by academic administrators. Therefore, the developed application can support more effective, efficient, and structured management of student report titles.

Abstrak

Meningkatnya jumlah laporan akhir mahasiswa setiap tahun menimbulkan tantangan bagi perguruan tinggi dalam menjaga keunikan dan kebaruan judul yang diajukan. Proses pengecekan judul yang masih dilakukan secara manual cenderung memerlukan waktu yang lama dan berpotensi menimbulkan ketidakonsistenan dalam proses validasi. Penelitian ini bertujuan untuk merancang dan membangun aplikasi Smart Title Checker berbasis website yang mampu mendeteksi tingkat kemiripan judul laporan mahasiswa menggunakan metode Cosine Similarity. Metode ini digunakan untuk mengukur tingkat kesamaan antarjudul sehingga dapat membantu proses validasi secara lebih objektif. Pengembangan sistem dilakukan melalui tahapan analisis kebutuhan, perancangan, implementasi, dan pengujian aplikasi. Hasil penelitian menunjukkan bahwa aplikasi mampu melakukan pengecekan kemiripan judul secara cepat dan akurat dengan menampilkan persentase similarity terhadap data judul yang tersimpan dalam basis data. Implementasi sistem juga membantu mahasiswa memperoleh informasi awal mengenai potensi kemiripan judul serta meningkatkan efisiensi proses validasi yang dilakukan oleh pengelola akademik. Dengan demikian, aplikasi yang dikembangkan dapat mendukung pengelolaan judul laporan mahasiswa secara lebih efektif, efisien, dan terstruktur.

This is an open access article under the CC-BY-SA license



1. Pendahuluan

Perguruan tinggi memiliki peran penting dalam menghasilkan sumber daya manusia yang kompeten, inovatif, dan mampu beradaptasi dengan perkembangan teknologi informasi. Salah satu bentuk implementasi kompetensi mahasiswa di lingkungan perguruan tinggi adalah penyusunan laporan akhir atau tugas akhir sebagai syarat penyelesaian studi. Laporan akhir tidak hanya menjadi bukti capaian akademik mahasiswa, tetapi juga berfungsi sebagai media pengembangan ilmu pengetahuan dan kontribusi terhadap penyelesaian berbagai permasalahan yang terjadi di masyarakat maupun dunia industri [1], [2].

Seiring dengan meningkatnya jumlah mahasiswa setiap tahun, jumlah laporan akhir yang dihasilkan juga mengalami peningkatan yang signifikan. Kondisi tersebut menyebabkan bertambahnya arsip judul laporan yang harus dikelola oleh program studi. Dalam praktiknya, sering ditemukan pengajuan judul yang memiliki kemiripan atau bahkan duplikasi dengan judul yang telah diajukan sebelumnya. Permasalahan ini dapat memengaruhi aspek kebaruan penelitian, kualitas karya ilmiah, serta efektivitas pengelolaan administrasi akademik. Oleh karena itu, proses validasi judul menjadi salah satu tahapan penting yang perlu dilakukan secara cermat untuk memastikan bahwa topik yang diajukan memiliki tingkat orisinalitas yang memadai [3], [4]. Pada banyak perguruan tinggi, proses pengecekan judul masih dilakukan secara manual dengan cara menelusuri arsip dokumen, repository, maupun daftar judul yang tersimpan pada file spreadsheet. Metode tersebut memerlukan waktu yang cukup lama dan sangat bergantung pada ketelitian petugas yang melakukan pemeriksaan. Selain itu, pencarian berdasarkan kata kunci sederhana sering kali belum mampu mendeteksi kemiripan antarjudul yang memiliki susunan kata berbeda tetapi membahas topik yang sama. Akibatnya, kemungkinan terjadinya duplikasi topik penelitian masih cukup tinggi meskipun proses verifikasi telah dilakukan [5], [6]. Perkembangan teknologi informasi memberikan peluang untuk meningkatkan efektivitas proses administrasi akademik melalui penerapan sistem berbasis web. Sistem berbasis web memungkinkan pengelolaan data secara terpusat, akses informasi secara real-time, serta mendukung proses pengambilan keputusan yang lebih cepat dan akurat. Pemanfaatan teknologi informasi dalam proses validasi judul laporan mahasiswa diharapkan mampu membantu program studi dalam mengurangi kesalahan manusia (*human error*) sekaligus meningkatkan kualitas layanan akademik [7], [8].

Dalam bidang text mining dan natural language processing (NLP), berbagai metode telah digunakan untuk mengukur tingkat kemiripan dokumen maupun teks. Penelitian yang dilakukan oleh Amalia et al. [9] menerapkan metode Cosine Similarity pada sistem ujian online otomatis untuk mengukur kesamaan jawaban. Rismayani et al. [10] memanfaatkan algoritma text mining dan Cosine Similarity untuk mendukung sistem aspirasi publik berbasis mobile. Penelitian lain yang dilakukan oleh Nasrullah [11] menunjukkan bahwa integrasi TF-IDF dan Cosine Similarity mampu meningkatkan kemampuan sistem dalam mendeteksi kemiripan judul penelitian. Selain itu, Rinjeni et al. [12] melakukan pencocokan judul artikel ilmiah menggunakan algoritma Cosine Similarity dan Jaccard Similarity untuk memperoleh tingkat kesamaan antarjudul secara lebih akurat. Meskipun berbagai penelitian tersebut telah berhasil menerapkan Cosine Similarity dalam pengukuran kemiripan teks, sebagian besar penelitian masih berfokus pada dokumen umum, artikel ilmiah, chatbot, maupun sistem pencarian informasi [13], [14], [15]. Penelitian-penelitian tersebut belum secara spesifik mengembangkan sistem terintegrasi yang dapat digunakan untuk membantu proses validasi judul laporan akhir mahasiswa pada lingkungan perguruan tinggi vokasi. Selain itu, sebagian penelitian hanya berfokus pada aspek algoritma tanpa mengintegrasikan mekanisme pengelolaan data judul, monitoring hasil pengecekan, serta dukungan pengambilan keputusan akademik secara langsung. Kondisi ini menunjukkan masih adanya kebutuhan penelitian terkait pengembangan

sistem pengecekan judul yang dapat digunakan secara praktis oleh mahasiswa, dosen, maupun pengelola program studi.

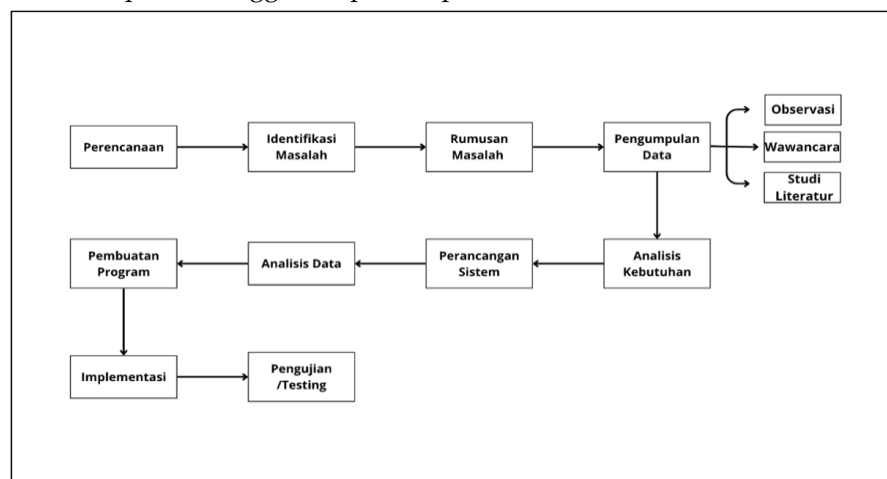
Metode Cosine Similarity dipilih dalam penelitian ini karena memiliki beberapa keunggulan dibandingkan pendekatan pencarian berbasis kata kunci sederhana maupun metode fuzzy matching. Pencarian berbasis kata kunci hanya mampu menemukan kesamaan berdasarkan kecocokan kata tertentu tanpa mempertimbangkan hubungan antar kata secara keseluruhan. Sementara itu, metode fuzzy matching lebih berorientasi pada kesamaan karakter dan kesalahan penulisan (typo), sehingga kurang optimal dalam mengukur kedekatan konteks antarjudul. Cosine Similarity bekerja dengan merepresentasikan dokumen dalam bentuk vektor dan menghitung tingkat kedekatan antarvektor, sehingga mampu memberikan pengukuran kemiripan yang lebih objektif meskipun terdapat perbedaan panjang kalimat atau variasi susunan kata [4], [10], [12].

Berdasarkan uraian tersebut, research gap dalam penelitian ini terletak pada belum tersedianya sistem Smart Title Checker berbasis website yang secara khusus dirancang untuk mendukung validasi judul laporan mahasiswa dengan memanfaatkan metode Cosine Similarity pada lingkungan perguruan tinggi vokasi. Oleh karena itu, penelitian ini bertujuan untuk merancang dan membangun aplikasi Smart Title Checker berbasis website yang mampu melakukan pengecekan kemiripan judul laporan mahasiswa secara otomatis. Sistem yang dikembangkan diharapkan dapat membantu mahasiswa memperoleh informasi awal mengenai potensi kemiripan judul, membantu dosen dan pengelola akademik dalam proses validasi, serta meningkatkan efektivitas dan objektivitas pengelolaan judul laporan akhir di Politeknik XYZ.

2. Metode Penelitian

2.1. Tahapan Penelitian

Tahapan penelitian merupakan serangkaian kegiatan yang dilakukan secara terencana, terstruktur, dan sistematis untuk mencapai tujuan penelitian yang telah ditetapkan (Dewi et al., 2021). Setiap tahapan disusun secara berurutan agar proses penelitian berjalan secara logis, mulai dari identifikasi permasalahan hingga diperolehnya hasil dan kesimpulan. Dengan adanya tahapan penelitian yang jelas, proses pemecahan masalah dapat dilakukan secara lebih terarah, efektif, dan sesuai dengan kaidah ilmiah. Gambar berikut menunjukkan tahapan penelitian yang dilakukan secara sistematis mulai dari tahap awal hingga tahap akhir penelitian

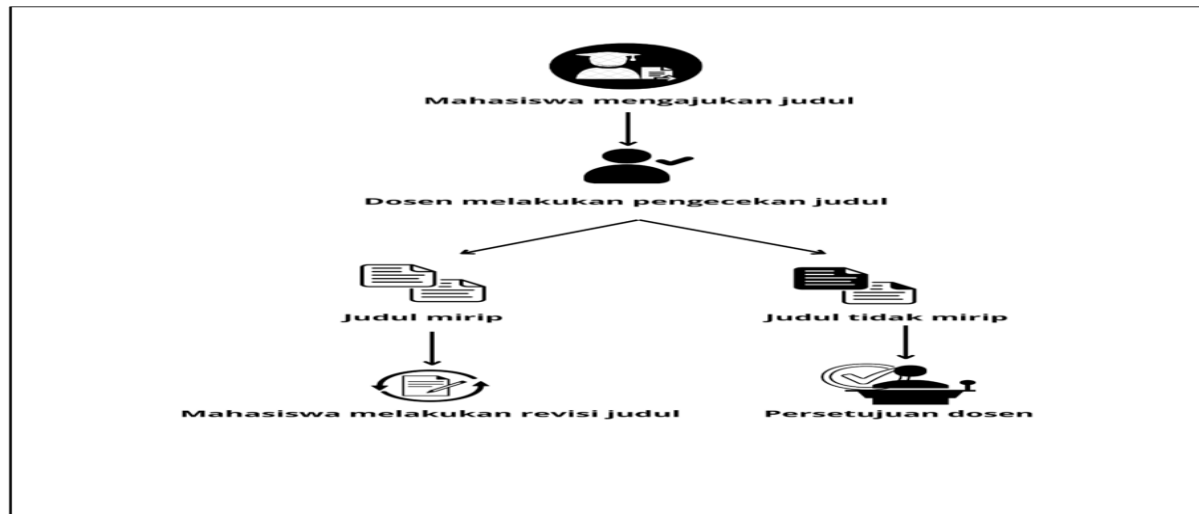


Gambar 1. Tahapan Penelitian

2.2. Alur Sistem Yang Berjalan

Proses pengecekan judul Laporan Akhir dan Tugas Akhir pada Program Studi Manajemen Informatika Politeknik Negeri Sriwijaya saat ini masih dilakukan secara semi-komputerisasi dengan

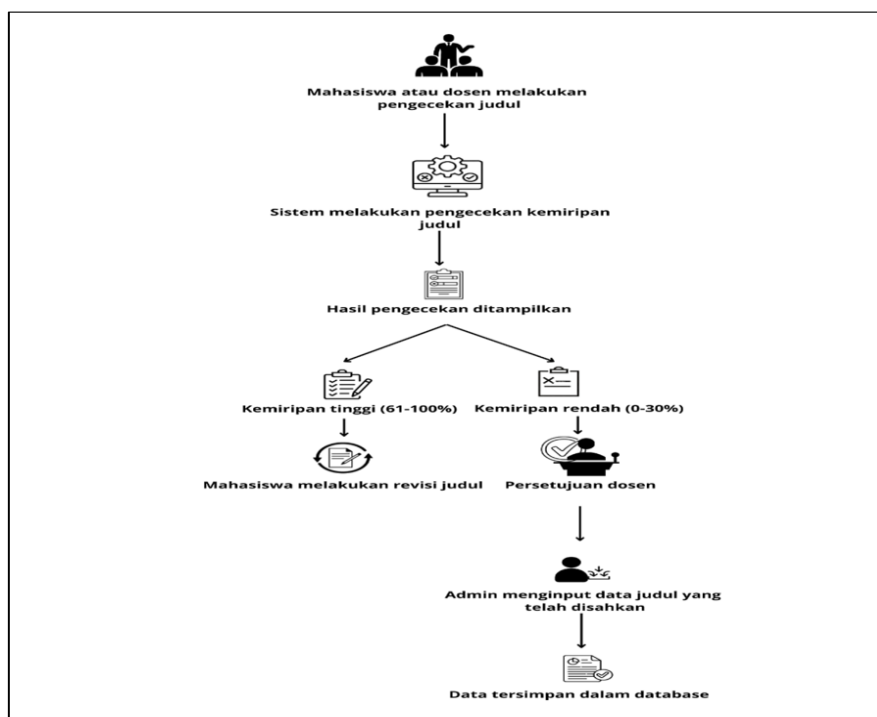
memanfaatkan arsip judul dalam bentuk *file* digital seperti *Microsoft Excel* dan *repository*. Meskipun telah menggunakan komputer, proses tersebut belum terintegrasi dalam sistem aplikasi sehingga pengecekan masih dilakukan secara terpisah dan bergantung pada ketelitian pengguna



Gambar 2. Alur Sistem Yang Berjalan

2.3. Sistem Yang Diusulkan

Manuskrip Sistem yang diusulkan pada proses pengecekan judul Laporan Akhir Mahasiswa Politeknik XYZ dirancang untuk mempermudah dan mengintegrasikan proses pengecekan judul. Sistem ini melibatkan mahasiswa, admin, dan dosen dalam proses pengecekan judul. Pada gambar 1.3 memperlihatkan alur sistem yang sedang diusulkan pada proses pengecekan judul Laporan Akhir Mahasiswa Politeknik XYZ.



Gambar 3. Alur Sistem Yang Berjalan

2.3. Metode Pengembangan sistem

Tabel Metode *Agile* merupakan metode pengembangan perangkat lunak yang banyak digunakan karena memiliki sifat fleksibel. Dalam prosesnya, pengembang dapat menyesuaikan kembali tahapan sebelumnya jika terjadi perubahan kebutuhan, sehingga sistem dapat terus disesuaikan dengan kondisi yang ada.

Agile Software Development adalah metode pengembangan sistem yang dirancang untuk menyesuaikan perubahan kebutuhan secara cepat. Selain itu, metode ini juga bertujuan menghasilkan perangkat lunak yang berkualitas dalam waktu yang relatif singkat melalui proses yang berulang dan terstruktur.

Berdasarkan kedua pendapat tersebut, dapat disimpulkan bahwa metode *Agile* merupakan pendekatan pengembangan perangkat lunak yang menekankan fleksibilitas dan kemampuan beradaptasi terhadap perubahan kebutuhan sistem. Metode ini memungkinkan proses pengembangan dilakukan secara bertahap dan berulang, sehingga hasil yang diperoleh dapat terus disesuaikan dengan kebutuhan pengguna. Selain itu, *Agile* juga membantu menghasilkan sistem dengan kualitas yang baik dalam waktu yang lebih cepat karena melibatkan proses evaluasi dan perbaikan secara berkelanjutan.

Metode *Agile* memiliki beberapa tahapan yang dilakukan secara berulang dalam proses pengembangan sistem. Tahapan tersebut dimulai dari perencanaan hingga pemeliharaan sistem agar dapat terus menyesuaikan dengan kebutuhan pengguna.

1. Perencanaan (*Planning*)

Tahap perencanaan dilakukan dengan mengumpulkan informasi kebutuhan sistem melalui observasi maupun wawancara dengan pengguna. Setelah itu, pengembang mulai menyusun rancangan sistem secara menyeluruh menggunakan alat bantu seperti UML, ERD, dan LRS agar sistem yang akan dibangun memiliki gambaran yang jelas.

2. Implementasi/Perancangan (*Implementation*)

Tahap implementasi merupakan proses pembuatan sistem sesuai dengan desain yang telah dirancang sebelumnya. Pada tahap ini, pengembang mulai melakukan *coding* menggunakan bahasa pemrograman serta *framework* yang dipilih untuk membangun aplikasi.

3. Pengujian (*Testing*)

Setelah sistem selesai dibuat, tahap selanjutnya adalah pengujian untuk memastikan sistem berjalan dengan baik. Pengujian dilakukan untuk menemukan kesalahan atau *bug* serta memastikan bahwa *input* dan *output* sistem sudah sesuai dengan yang diharapkan.

4. Dokumentasi (*Documentation*)

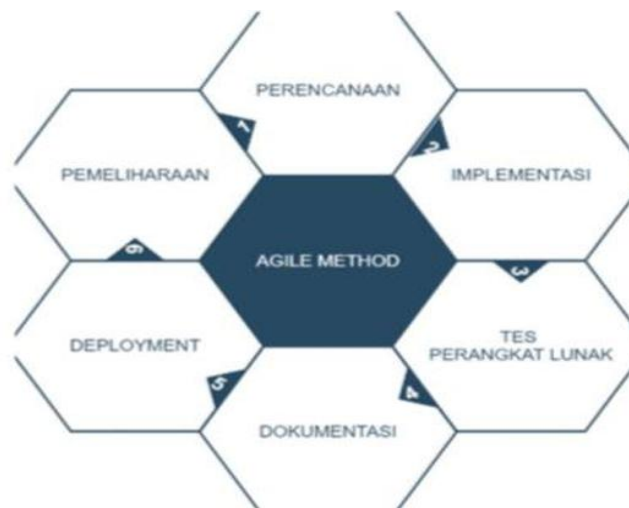
Tahap dokumentasi dilakukan dengan mencatat seluruh proses pengembangan sistem, termasuk modul dan fungsi yang ada. Dokumentasi ini berguna sebagai referensi untuk pengembangan sistem di masa mendatang.

5. Penyebaran (*Deployment*)

Pada tahap ini, sistem yang telah selesai dikembangkan mulai diterapkan agar dapat digunakan oleh pengguna secara langsung.

6. Pemeliharaan (*Maintenance*)

Tahap pemeliharaan dilakukan dengan memantau sistem secara berkala untuk memastikan sistem tetap berjalan dengan baik serta memperbaiki jika ditemukan kesalahan atau celah pada sistem.



Gambar 4. Alur Sistem Yang Berjalan

2.4. Metode Perhitungan *Cosine Similiarty*

Cosine similarity merupakan teknik untuk mengukur tingkat kemiripan di antara dua vektor berdimensi n melalui perhitungan nilai kosinus dari sudut yang terbentuk (Lindang et al., 2022). Dalam bidang *text mining*, pendekatan ini secara luas diimplementasikan untuk melakukan komparasi antar dokumen. Rumus cosine similiarty dapat dilihat dibawah ini:

$$Similarity(x, y) = \cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|}$$

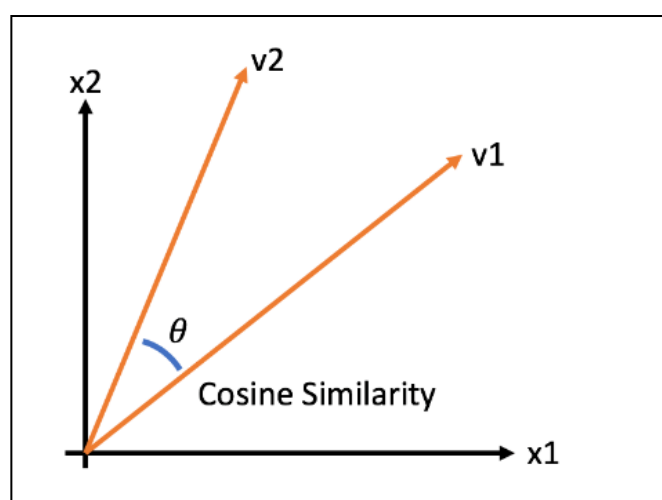
Dimana:

$x \cdot y$: vector dot product dari x dan y , dihitung dengan $\sum_{k=1}^n x_k y_k$;

$\|x\|$: panjang vektor x , dihitung dengan $\sqrt{\sum_{k=1}^n x_k^2}$;

$\|y\|$: panjang vektor y , dihitung dengan $\sqrt{\sum_{k=1}^n y_k^2}$

Cosine Similarity merupakan metode yang digunakan untuk mengukur tingkat kesamaan antar dokumen (Basuki et al., 2023). Dalam penerapannya, metode ini memanfaatkan representasi himpunan atau vektor, dengan perhitungan yang didasarkan pada rumus *cosine similarity*.



Gambar 5. Alur Sistem Yang Berjalan

Berdasarkan kedua pendapat tersebut, dapat disimpulkan bahwa *cosine similarity* merupakan metode yang digunakan untuk mengukur tingkat kemiripan antar dokumen dengan

merepresentasikan data dalam bentuk vektor. Metode ini bekerja dengan menghitung nilai kosinus dari sudut yang terbentuk antara dua vektor, sehingga mampu menunjukkan tingkat kesamaan berdasarkan kedekatan arah vektor tersebut. *Cosine similarity* banyak digunakan dalam bidang *text mining* karena efektif dalam membandingkan dokumen meskipun memiliki panjang teks yang berbeda.

2.5. Dataset Penelitian dan Teknik Pengujian

Dataset yang digunakan dalam penelitian ini berupa kumpulan arsip judul Laporan Akhir mahasiswa Program Studi Manajemen Informatika Politeknik XYZ yang tersimpan pada basis data akademik. Data yang digunakan merupakan judul laporan yang telah disetujui oleh program studi dan terdokumentasi secara resmi dalam repository institusi. Dataset dipilih karena merepresentasikan kondisi nyata yang dihadapi oleh pengelola akademik dalam proses validasi judul laporan mahasiswa.

Data yang digunakan terdiri dari 250 judul laporan mahasiswa yang berasal dari periode akademik 2021–2025. Setiap judul memiliki karakteristik panjang kalimat antara 5 hingga 15 kata dengan topik penelitian yang beragam, seperti sistem informasi, aplikasi berbasis web, aplikasi mobile, data mining, kecerdasan buatan, dan teknologi informasi. Keberagaman topik tersebut bertujuan untuk menguji kemampuan metode Cosine Similarity dalam mendeteksi tingkat kemiripan antarjudul pada berbagai bidang kajian.

Tabel 1. Karakteristik Dataset

No	Karakteristik Dataset	Keterangan
1	Jenis Data	Judul Laporan Akhir Mahasiswa
2	Sumber Data	Repository dan Arsip Akademik Politeknik XYZ
3	Jumlah Dataset	250 Judul
4	Periode Data	2021–2025
5	Panjang Judul	5–15 Kata
6	Bidang Kajian	Sistem Informasi, Web, Mobile, Data Mining, AI
7	Format Penyimpanan	Database MySQL
8	Metode Analisis	Cosine Similarity

Untuk mengevaluasi kinerja sistem, dilakukan pengujian menggunakan 30 data judul sebagai data uji yang dibandingkan dengan seluruh data yang tersimpan pada basis data. Setiap judul diproses melalui tahapan preprocessing, pembentukan vektor kata, serta perhitungan nilai Cosine Similarity untuk memperoleh tingkat kemiripan terhadap judul yang telah ada sebelumnya.

Tabel 2. Pengujian Sistem

No	Parameter Pengujian	Nilai
1	Jumlah Data Uji	30 Judul
2	Dataset Pembanding	250 Judul
3	Metode Pengujian	Black Box Testing
4	Metode Similarity	Cosine Similarity
5	Output Sistem	Persentase Similarity
6	Ambang Batas (Threshold)	≥ 80% Tidak Valid
7	Kategori Valid	< 80%
8	Jumlah Pengujian	30 Kali Pengujian

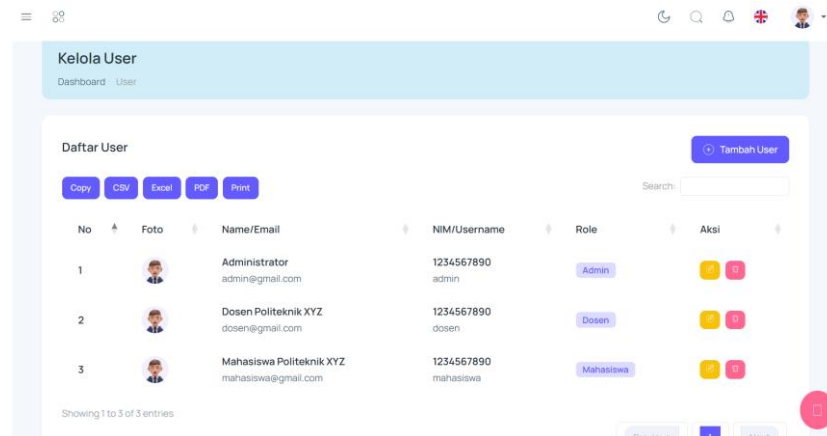
Selain melakukan pengujian fungsional menggunakan metode Black Box Testing, penelitian ini juga mengevaluasi kemampuan sistem dalam mengidentifikasi kemiripan judul berdasarkan nilai similarity yang dihasilkan. Hasil pengujian menunjukkan bahwa sistem mampu mendeteksi judul yang identik dengan tingkat kemiripan mencapai 100% serta mampu membedakan judul yang memiliki tingkat kemiripan rendah di bawah 40%. Dengan demikian, sistem yang dikembangkan dapat digunakan sebagai alat bantu dalam proses validasi judul laporan mahasiswa secara lebih objektif, cepat, dan terukur.

3. Hasil dan Pembahasan

Berdasarkan hasil dari analisis dan implementasi yang dilakukan sistem yang dibuat terdiri dari beberapa menu diantaranya kelola user, input data judul laporan serta prosesnya dan monitoring hasil pengecekan, adapun penjelasannya adalah sebagai berikut:

3.1. Halaman Kelola User

Pada halaman kelola user, sistem menyediakan fitur untuk mengelola data pengguna yang terdiri dari administrator, dosen, dan mahasiswa.

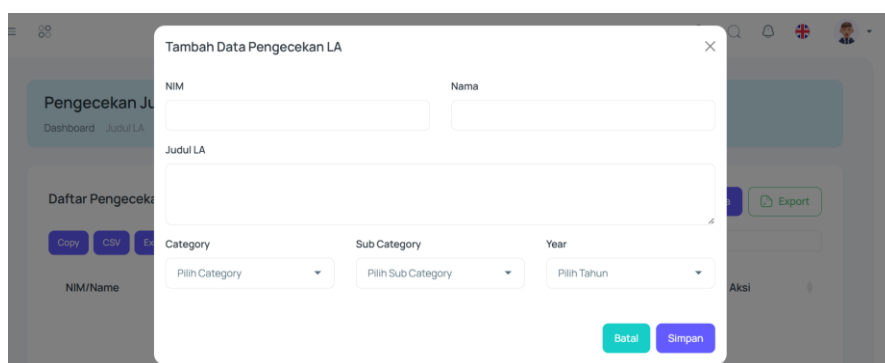


Gambar 6. Kelola User

Fitur yang tersedia pada halaman ini meliputi Menambahkan data pengguna baru. Mengubah data pengguna. Menghapus data pengguna. Menentukan hak akses berdasarkan peran pengguna. Berdasarkan implementasi yang dilakukan, sistem berhasil menampilkan informasi pengguna berupa nama, email, NIM/username, dan role pengguna secara terstruktur sehingga memudahkan pengelolaan data pengguna.

3.2. Halaman Input Data Judul Laporan

Halaman ini merupakan bagian dimana user menginputkan data judul laporan akhir mahasiswa yang akan diajukan ke sistem.



Gambar 7. Input Data Judul Laporan

Pada halaman ini mahasiswa harus menginputkan nim, nama serta judul laporan akhir yang akan diajukan untuk diperiksa, pada saat sistem mengecek data sistem akan menyimpan data judul tersebut ke database dan akan menampilkan notifikasi ketika proses penyimpanan berhasil.

3.3. Halaman Hasil Pengecekan Judul

Halaman pengecekan merupakan fitur utama dalam aplikasi. Sistem melakukan perhitungan kemiripan judul menggunakan metode Cosine Similarity dengan membandingkan judul yang diinput dengan data judul yang tersimpan pada database.

ID	Judul	Deskripsi	Persentase Kemiripan	Status
321312313123	budi test	Aplikasi Penentuan Pemasaran Terbanyak Konsumen dalam Pembelian Rumah dengan Metode Weighted Product (WP) di PT Cipta Bangun Masi	100%	Pending
321312313124	muhammad bambang	Penerapan Metode Single Moving Average (SMA) dalam Aplikasi Prediksi Pengadaan Barang Display pada PT Sasa Inti	100%	Pending
61940830322	Indah Apriani	Aplikasi Penentuan Pemasaran Terbanyak Konsumen dalam Pembelian Rumah dengan Metode Weighted Product (WP) di PT Cipta Bangun Masi	30.43%	Pending
61940830323	Lemi Yensu	Aplikasi Pengukuran Indeks Kepuasan Pelanggan Terhadap Kinerja Layanan Produk Lion Parcel Post Seberang pada PT Lion Express Menggunakan k-Medoids Clustering	29.81%	Pending

Gambar 8. Hasil Pengecekan Judul

Berdasarkan hasil pengecekan diatas yang dilakukan oleh sistem terdapat beberapa hasil similarity sebagai berikut :

Tabel 3. Hasil pengecekan

No	Judul yang Dicek	Persentase Similarity	Status
1	Data pertama	100%	Tidak Valid
2	Data kedua	100%	Tidak Valid
3	Data ketiga	30.43%	Valid
4	Data keempat	29.81%	Valid

Dari hasil tersebut terlihat bahwa sistem mampu mendeteksi judul yang identik maupun yang memiliki tingkat kemiripan rendah.

3.4. Pembahasan Metode Cosine Similarity

Metode Cosine Similarity digunakan untuk menghitung tingkat kemiripan antar judul dengan membandingkan representasi vektor dari setiap kata pada dokumen.

Secara umum perhitungan similarity dilakukan menggunakan rumus berikut:

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|}$$

Keterangan:

A = vektor judul input

B = vektor judul pembanding

A · B = hasil perkalian dot product

|| A || dan || B || = panjang vektor

Nilai similarity berada pada rentang 0-1, dengan interpretasi:

0 = tidak memiliki kemiripan

1 = identik

Pada implementasi sistem:

Nilai 100% menunjukkan judul identik dengan data yang tersimpan.

Nilai 30.43% menunjukkan adanya beberapa kata yang sama namun konteks berbeda.

Nilai 29.81% menunjukkan tingkat kemiripan rendah.

Hasil tersebut menunjukkan bahwa metode *Cosine Similarity* mampu memberikan pengukuran yang objektif terhadap tingkat kemiripan judul laporan mahasiswa

3.5. Analisis Hasil dan Evaluasi Sistem

Berdasarkan hasil pengujian yang telah dilakukan, aplikasi Smart Title Checker mampu mengidentifikasi tingkat kemiripan judul laporan mahasiswa secara otomatis dengan memanfaatkan metode *Cosine Similarity*. Hasil pengujian menunjukkan bahwa judul yang identik menghasilkan nilai similarity sebesar 100%, sedangkan judul yang memiliki perbedaan konteks menghasilkan nilai similarity yang lebih rendah, yaitu sebesar 30,43% dan 29,81%. Hasil tersebut menunjukkan bahwa metode *Cosine Similarity* mampu membedakan antara judul yang memiliki tingkat kesamaan tinggi dan judul yang hanya memiliki beberapa kata yang serupa.

Kemampuan tersebut menjadi salah satu keunggulan *Cosine Similarity* dibandingkan metode pencarian berbasis kata kunci sederhana. Pada metode pencarian kata kunci, sistem hanya melakukan pencocokan berdasarkan kemunculan kata tertentu sehingga sering menghasilkan kesalahan identifikasi ketika terdapat variasi susunan kata. Sebaliknya, *Cosine Similarity* melakukan pengukuran berdasarkan representasi vektor sehingga hubungan antar kata dapat dianalisis secara lebih objektif. Dengan pendekatan tersebut, tingkat kemiripan dapat dihitung secara kuantitatif dan menghasilkan nilai yang lebih representatif terhadap kedekatan isi judul [4], [10].

Selain itu, metode *Cosine Similarity* juga memiliki keunggulan dibandingkan pendekatan fuzzy matching yang umumnya digunakan untuk mendeteksi kesalahan penulisan atau kemiripan karakter. Fuzzy matching lebih efektif untuk mengidentifikasi typo, namun kurang optimal dalam membandingkan dokumen atau judul yang memiliki susunan kata berbeda tetapi membahas topik yang sama. Oleh karena itu, penggunaan *Cosine Similarity* dinilai lebih sesuai untuk proses validasi judul laporan mahasiswa yang membutuhkan pengukuran kemiripan berdasarkan keseluruhan isi teks [5], [11].

Dari sisi performa sistem, pengujian terhadap 250 data judul yang tersimpan dalam basis data menunjukkan bahwa proses pencarian dan perhitungan similarity dapat dilakukan dalam waktu yang relatif singkat. Sistem mampu menampilkan hasil pengecekan secara real-time setelah pengguna memasukkan judul yang akan diperiksa. Hal ini menunjukkan bahwa metode yang diterapkan masih mampu bekerja dengan baik pada jumlah data yang digunakan dalam penelitian. Seiring dengan bertambahnya jumlah data di masa mendatang, optimasi proses indexing dan preprocessing dapat diterapkan untuk menjaga performa sistem tetap optimal.

Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh Nasrullah [4] yang menyatakan bahwa *Cosine Similarity* efektif digunakan untuk mendeteksi kemiripan judul penelitian. Temuan ini juga mendukung penelitian Rinjeni et al. [5] yang menunjukkan bahwa metode *Cosine Similarity* mampu memberikan tingkat pengukuran kemiripan yang lebih objektif dibandingkan pendekatan pencocokan teks secara sederhana. Dengan demikian, aplikasi Smart Title Checker yang dikembangkan tidak hanya mampu membantu proses administrasi akademik, tetapi juga memberikan dukungan pengambilan keputusan yang lebih terukur dalam proses validasi judul laporan mahasiswa

Tabel 4. Perbandingan Metode Pendeteksian Kemiripan Judul

Aspek	<i>Keyword Matching</i>	<i>Fuzzy Matching</i>	<i>Cosine Similarity</i>
Pencocokan Kata	Ya	Ya	Ya
Deteksi Typo	Tidak	Ya	Terbatas
Analisis Kemiripan Konteks	Tidak	Tidak	Ya
Nilai Similarity Kuantitatif	Tidak	Ya	Ya

Cocok untuk Validasi Judul	Cukup	Baik	Sangat Baik
----------------------------	-------	------	-------------

4. Kesimpulan

Berdasarkan hasil penelitian, aplikasi Smart Title Checker berbasis website berhasil dikembangkan untuk membantu proses validasi judul laporan mahasiswa secara otomatis menggunakan metode Cosine Similarity. Sistem mampu menghitung tingkat kemiripan antarjudul dalam bentuk persentase, mendeteksi judul identik dengan similarity 100%, serta membedakan judul dengan tingkat kemiripan rendah di bawah 40%. Seluruh fitur utama, seperti pengelolaan pengguna, input judul, pengecekan kemiripan, dan monitoring hasil, berjalan sesuai kebutuhan. Aplikasi ini berkontribusi dalam meningkatkan efektivitas dan efisiensi proses administrasi akademik dengan membantu mahasiswa dan pengelola akademik melakukan verifikasi judul secara lebih cepat dan terukur. Namun, penelitian masih memiliki keterbatasan pada jumlah dan cakupan dataset yang hanya berasal dari satu program studi dan institusi, serta belum membandingkan metode Cosine Similarity dengan algoritma lain. Penelitian selanjutnya disarankan menggunakan dataset yang lebih luas, mengintegrasikan sistem dengan repository akademik secara real-time, serta menerapkan metode hybrid berbasis text mining atau Natural Language Processing untuk meningkatkan akurasi dan skalabilitas deteksi kemiripan judul.

Daftar Pustaka

- [1] E. L. Amalia, A. J. Jumadi, I. A. Mashudi, and D. W. Wibowo, "Analisis metode cosine similarity pada aplikasi ujian online otomatis," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 2, pp. 343–350, 2021.
- [2] R. Rismayani, H. Sy, T. Darwansyah, and I. Mansyur, "Implementasi algoritma text mining dan cosine similarity untuk desain sistem aspirasi publik berbasis mobile," *Komputika: Jurnal Sistem Komputer*, vol. 11, no. 2, pp. 169–176, 2022.
- [3] D. G. Purnama et al., "Application of natural language processing using cosine-similarity algorithm in making chatbot information," in *2022 7th International Workshop on Big Data and Information Security*, pp. 1–6, 2022.
- [4] H. Nasrullah, "Integrasi TF-IDF dan algoritma cosine similarity untuk deteksi tingkat kemiripan judul penelitian," *Information Technology Education Journal*, vol. 3, no. 3, pp. 113–118, 2024.
- [5] T. P. Rinjeni, A. Indriawan, and N. A. Rakhmawati, "Matching scientific article titles using cosine similarity and Jaccard similarity algorithm," *Procedia Computer Science*, vol. 234, pp. 553–560, 2024.
- [6] D. Sunandar and A. Muiz, "Thesis title similarity detection system using Levenshtein distance and cosine similarity," *bit-Tech*, vol. 8, no. 1, pp. 1–10, 2025.
- [7] O. A. Resta, A. Aditya, and F. E. Purwiantono, "Plagiarism detection in students' theses using the cosine similarity method," *Sinkron*, vol. 5, no. 2, pp. 1–10, 2025.
- [8] M. F. Ansyori and A. H. Mujianto, "Penerapan natural language processing dengan metode cosine similarity pada sistem e-monev," *Journal Software, Hardware and Information Technology*, vol. 5, no. 2, pp. 1–12, 2025.
- [9] H. Sari et al., "Penerapan algoritma text mining dan TF-IDF untuk pengelompokan topik skripsi pada aplikasi repository," *Terapan Informatika Nusantara*, vol. 2, no. 7, pp. 414–432, 2021.
- [10] A. Sanjaya and S. D. Sasongko, "Uji kemiripan kalimat menggunakan preprocessing dan cosine similarity dalam bahasa Indonesia," *Jurnal Ilmiah Nero*, vol. 7, no. 2, pp. 95–104, 2022.
- [11] R. Al Rasyid and D. H. U. Ningsih, "Penerapan algoritma TF-IDF dan cosine similarity untuk query pencarian pada dataset destinasi wisata," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 8, no. 1, pp. 170–178, 2024.
- [12] R. Huang, "Improved content recommendation algorithm integrating semantic information," *Journal of Big Data*, vol. 10, no. 84, pp. 1–20, 2023.
- [13] G. E. I. Kambey et al., "Penerapan clustering pada aplikasi pendeteksi kemiripan dokumen teks bahasa Indonesia," *Jurnal Informatika*, vol. 15, no. 2, pp. 75–82, 2020.
- [14] A. Fauzi, B. Rahman, and S. Hidayat, "Implementasi sistem informasi berbasis web untuk pengelolaan data akademik perguruan tinggi," *Jurnal Sistem Informasi dan Teknologi Informasi*, vol. 6, no. 1, pp. 45–54, 2021.

- [15] M. Yusuf, R. Pratama, and N. Saputra, "Pengembangan aplikasi repository karya ilmiah mahasiswa berbasis web menggunakan metode prototype," *Jurnal Media Informatika*, vol. 7, no. 2, pp. 88-97, 2023.