

SENASTIKA Universitas Malikussaleh

IMPLEMENTASI SISTEM REKOMENDASI CONTENT-BASED FILTERING UNTUK REPOSITORI SKRIPSI TEKNIK INFORMATIKA MENGGUNAKAN COSINE SIMILARITY DAN JACCARD INDEX

Abil Khairi^{*1}, Wahyu Fuadi², Yesy Afrillia^{*3}

^{1,2,3}Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh, Indonesia

Email: abil.200170168@mhs.unimal.ac.id, yesy.afrillia@unimal.ac.id, wahyu.fuadi@unimal.ac.id

Abstrak

Jurusan Teknik Informatika Universitas Malikussaleh sering menghadapi masalah dalam membantu mahasiswa menemukan topik skripsi yang sesuai dengan minat dan keterampilan akademik mereka. Masalah ini juga diperburuk oleh kurangnya repositori skripsi yang efektif, yang menyebabkan duplikasi judul dan ketidakefisienan dalam pemilihan topik. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi berbasis *content-based filtering* yang memanfaatkan teknologi *Cosine Similarity* dan *Jaccard Index* sebagai metode untuk mengidentifikasi kesamaan antar dokumen skripsi. Sistem ini dirancang untuk membantu mahasiswa memilih topik skripsi yang relevan dengan minat mereka serta meningkatkan efisiensi pengelolaan repositori skripsi oleh pihak jurusan. Metode yang digunakan mencakup pengumpulan data skripsi dari repositori jurusan dan *web scraping*, diikuti dengan prapemrosesan data menggunakan teknik tokenisasi, penghilangan *stop-words*, dan *stemming*. Evaluasi sistem dilakukan dengan mengukur presisi, *recall*, dan *F1-score*, dengan hasil akhir menunjukkan bahwa sistem memiliki presisi sebesar 80% dan *recall* sebesar 60%. Hasil penelitian ini menunjukkan bahwa sistem mampu meningkatkan proses pemilihan topik skripsi, namun masih diperlukan peningkatan dalam akurasi dan variasi rekomendasi.

Kata Kunci: *Content-Based Filtering, Cosine Similarity, Jaccard Index, Repositori Skripsi, Sistem Rekomendasi, Universitas Malikussaleh.*

Abstract

The Informatics Engineering Department at Universitas Malikussaleh often faces challenges in assisting students to find thesis topics that align with their interests and academic skills. This issue is exacerbated by the lack of an effective thesis repository, leading to title duplication and inefficiency in topic selection. This study aims to develop a recommendation system based on *content-based filtering*, utilizing *Cosine Similarity* and *Jaccard Index* as methods to identify similarities between thesis documents. The system is designed to help students choose thesis topics relevant to their interests while improving the department's efficiency in managing the thesis repository. The method includes collecting thesis data from the department's repository and web scraping, followed by data preprocessing techniques such as tokenization, stop-word removal, and stemming. The system evaluation measured precision, recall, and F1-score, with results showing a precision of 80% and a recall of 60%. This research demonstrates that the system improves the thesis topic selection process, although further improvements in accuracy and recommendation variety are still needed.

Keywords: *Content-Based Filtering, Cosine Similarity, Jaccard Index, Theses Repository, Recommendation System, Universitas Malikussaleh.*

1. PENDAHULUAN

Skripsi menjadi syarat kelulusan yang mencerminkan kemampuan penelitian mahasiswa, sebagaimana diamanatkan oleh Tri Dharma Perguruan Tinggi. Di Jurusan Teknik Informatika Universitas Malikussaleh, mahasiswa diwajibkan untuk menyusun skripsi sebagai bukti kemampuan akademik mereka dalam bidang ilmu yang mereka pelajari. Namun, memilih topik skripsi sering kali menghadapi berbagai tantangan.

Mahasiswa Teknik Informatika di Universitas Malikussaleh sering mengalami kesulitan dalam menemukan topik skripsi yang sesuai dengan minat, keterampilan, dan kebutuhan akademik. Masalah ini diperburuk oleh tidak adanya repositori skripsi yang efektif untuk menyimpan, mengelola, dan merekomendasikan skripsi yang sudah ada. Akibatnya, mengalami kesulitan dalam menemukan referensi skripsi yang relevan. Selain itu, pengelolaan

topik skripsi secara manual di Jurusan Teknik Informatika sering menyebabkan masalah seperti duplikasi judul dan ketidakefisienan dalam pemilihan topik.

Untuk mengatasi tantangan ini, diperlukan sebuah sistem yang dapat membantu mahasiswa dalam menemukan topik skripsi yang sesuai secara lebih efektif. Sistem ini seharusnya memberikan rekomendasi topik yang dipersonalisasi berdasarkan minat, keterampilan, dan latar belakang akademis mahasiswa, menyediakan repositori lengkap dari skripsi yang ada sebagai referensi, serta menyederhanakan proses pemilihan topik skripsi agar terhindar dari duplikasi.

Penelitian terdahulu telah menunjukkan pentingnya sistem rekomendasi dalam konteks akademik, sebagaimana yang dijelaskan dalam berbagai studi. Kusuma dan Musdholifah (2021) dalam penelitian mereka tentang "Sistem Rekomendasi untuk Topik Skripsi Menggunakan *Content-based filtering*" berhasil mencapai tingkat keberhasilan sebesar 83% dengan menerapkan metode *Content-based filtering* untuk menganalisis elemen-elemen seperti judul, abstrak, dan kata kunci pada dokumen skripsi, sehingga mampu memberikan rekomendasi yang sesuai dengan preferensi individu. Di sisi lain, penelitian yang dilakukan oleh Lukas Kurniawan et al. (2023) terkait "Sistem rekomendasi Topik Skripsi Program Studi Informatika" berhasil mengembangkan sistem rekomendasi topik skripsi yang efisien dan akurat dengan menggunakan kombinasi algoritma K-Means++, *Cosine similarity*, dan LDA Gibbs Sampling. Sementara itu, Liga Mayola et al. (2024) dalam studinya tentang "Algoritma *Jaccard Similarity* untuk Deteksi Kemiripan Judul Disertasi dengan Pendekatan Variasi Stop Word Removal" menunjukkan bahwa penggunaan algoritma *Jaccard Similarity* dengan teknik Stop Word Removal Versi Dua (SWR2) lebih akurat dalam mendeteksi kemiripan judul disertasi dibandingkan dengan SWR1. Indriani et al. (2024) memanfaatkan Jaccard Index dan N-Gram dalam pengembangan aplikasi koreksi kata berbahasa Indonesia, di mana penerapan Jaccard Index terbukti meningkatkan akurasi dalam mendeteksi dan memperbaiki kesalahan penulisan, yang relevan untuk pengembangan sistem rekomendasi berbasis analisis kesamaan dokumen.

Penelitian-penelitian tersebut menjadi landasan bagi penulis untuk mengusulkan pengembangan sistem rekomendasi content-based untuk judul skripsi. Sistem ini direncanakan untuk dikembangkan sebagai situs web repositori skripsi khusus untuk Jurusan Teknik Informatika Universitas Malikussaleh. Pendekatan content-based dipilih karena kemampuannya untuk menganalisis elemen-elemen penting dari dokumen skripsi, seperti judul, abstrak, dan kata kunci, guna menghasilkan rekomendasi yang relevan dan sesuai dengan preferensi individu. Dalam implementasinya, sistem ini akan menggunakan teknik Similarity Measure, termasuk *Cosine similarity* dan Jaccard Index, yang dipilih karena keefektifannya dalam membandingkan teks dokumen dan menghasilkan rekomendasi yang akurat berdasarkan kesamaan konten.

Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi berbasis *Content-based filtering* yang dapat digunakan oleh mahasiswa di Jurusan Teknik Informatika Universitas Malikussaleh. Tujuan utama dari sistem ini adalah untuk:

1. Membantu mahasiswa dalam menemukan topik skripsi yang sesuai dengan minat dan keterampilan mereka.
2. Mengurangi risiko terjadinya duplikasi topik skripsi.
3. Meningkatkan efisiensi dalam proses pemilihan dan persetujuan topik skripsi.
4. Memfasilitasi pengelolaan repositori skripsi oleh pihak jurusan, sehingga mempermudah akses terhadap referensi penelitian yang ada.

Pengembangan sistem ini diharapkan akan mempermudah mahasiswa dalam menemukan topik skripsi yang sesuai dengan minat dan kemampuan mereka, serta relevan dengan kebutuhan akademik. Sistem ini juga diharapkan dapat meningkatkan efisiensi dalam proses pemilihan topik skripsi, mengurangi risiko duplikasi topik, dan membantu jurusan dalam mengelola dan mengorganisir skripsi yang ada.

2. METODE PENELITIAN

Penelitian ini berfokus pada pengembangan sistem rekomendasi berbasis content-based filtering untuk repositori skripsi di Jurusan Teknik Informatika Universitas Malikussaleh. Permasalahan utama yang dihadapi adalah bagaimana mengidentifikasi dan merekomendasikan skripsi yang relevan berdasarkan konten untuk membantu mahasiswa dalam pemilihan topik dan referensi skripsi.

Formulasi masalah dapat dinyatakan secara matematis sebagai berikut:

Menemukan subset dokumen $R \subseteq D$, dimana $R = \{r_1, r_2, \dots, r_k\}$ adalah k dokumen yang paling relevan dengan q , sehingga:

$$\forall r_i \in D \setminus R : \text{sim}(q, r_i) \geq \text{sim}(q, d_j)$$

Metode yang diusulkan menggunakan kombinasi *Cosine Similarity* dan *Jaccard Index* untuk menghitung kemiripan antar dokumen:

$$\text{sim}(q, d_i) = \alpha \cdot \text{sim}_{\text{cos}}(q, d_i) + (1 - \alpha) \cdot \text{sim}_{\text{jac}}(q, d_i)$$

dimana:

$$\begin{aligned} \text{sim}_{\cos}(q, d_i) &= \frac{q \cdot d_i}{\|q\| \|d_i\|} \\ \text{sim}_{\text{jac}}(q, d_i) &= \frac{|q \cap d_i|}{|q \cup d_i|} \end{aligned}$$

α adalah parameter bobot dengan $0 \leq \alpha \leq 1$

Untuk menyelesaikan permasalahan ini, penelitian dilakukan melalui beberapa tahapan. Pertama, dilakukan tinjauan mendalam terhadap literatur terkait sistem rekomendasi, content-based filtering, Cosine Similarity, Jaccard Index, dan teknologi web yang relevan. State of the art dalam bidang ini mencakup penggunaan teknik deep learning seperti Doc2Vec untuk representasi dokumen dan penggunaan metrik kemiripan hibrid untuk meningkatkan akurasi rekomendasi.

Selanjutnya, pengumpulan data dilakukan dari dua sumber utama: repositori Jurusan Teknik Informatika (241 data skripsi) dan hasil web scraping dari <https://rama.unimal.ac.id/> (lebih dari 200 data skripsi). Data yang terkumpul kemudian melalui tahap prapemrosesan yang meliputi tokenisasi, penghilangan stop-words, stemming, dan lemmatization menggunakan library NLTK dan Sastrawi untuk Bahasa Indonesia.

Sistem dikembangkan menggunakan metode Waterfall, mencakup analisis kebutuhan, desain sistem, implementasi, pengujian, dan pemeliharaan. Implementasi backend menggunakan Python dengan FastAPI, sementara frontend dikembangkan dengan Next.js. Visual Studio Code digunakan sebagai lingkungan pengembangan utama untuk memastikan konsistensi dan efisiensi dalam proses pengembangan.

Cara kerja sistem dimulai dengan pengguna memasukkan indeks skripsi sebagai query. Sistem kemudian menganalisis konten skripsi yang diinput, mencari konten serupa dalam database menggunakan algoritma kemiripan yang telah didefinisikan, dan akhirnya menampilkan output berupa skor kemiripan dan indeks skripsi yang serupa.

Evaluasi sistem dilakukan menggunakan metrik F1-score, precision, dan recall. Tujuan evaluasi adalah untuk mengukur akurasi dan efektivitas sistem dalam mengidentifikasi dan menyajikan skripsi dengan konten serupa, serta memvalidasi keandalan sistem dalam membantu mahasiswa menemukan referensi yang relevan untuk penelitian mereka.

Pendekatan ini dipilih untuk menghasilkan sistem rekomendasi yang efektif, menggabungkan kekuatan dari pengolahan bahasa alami, algoritma kemiripan, dan analisis konten untuk membantu mahasiswa dalam memilih topik skripsi yang sesuai dengan minat dan kemampuan mereka. Dengan memformulasikan masalah secara matematis dan menerapkan metode yang terstruktur, penelitian ini bertujuan untuk memberikan solusi yang dapat diukur dan divalidasi secara objektif.

3. HASIL DAN PEMBAHASAN

Penelitian ini dilaksanakan dengan tujuan utama untuk mengembangkan sistem rekomendasi berbasis *Content-based filtering* untuk repositori skripsi di Jurusan Teknik Informatika Universitas Malikussaleh. Untuk mencapai tujuan tersebut, serangkaian langkah strategis telah diambil guna memastikan keberhasilan sistem dalam memberikan rekomendasi yang relevan bagi mahasiswa.

Tahap awal dimulai dengan pengumpulan data yang dilakukan melalui dua sumber utama. Pertama, repositori skripsi jurusan yang menyediakan informasi skripsi dari mahasiswa selama periode 2014 hingga 2024. Pengumpulan data ini dilakukan secara manual, mencakup judul, abstrak, dan kata kunci dari skripsi yang telah disusun. Kedua, data tambahan diperoleh melalui proses web scraping dari situs resmi Universitas Malikussaleh, yang juga menyimpan informasi skripsi mahasiswa.

Setelah data terkumpul, langkah selanjutnya adalah prapemrosesan data, yang meliputi tokenisasi untuk memecah teks menjadi kata-kata, penghapusan stop words untuk menghilangkan kata-kata yang tidak signifikan, dan stemming untuk mengembalikan kata-kata ke bentuk dasarnya. Proses prapemrosesan ini penting untuk memastikan bahwa analisis data yang dilakukan dapat memberikan hasil yang akurat dan relevan.

Sistem rekomendasi yang dikembangkan diuji untuk menilai kinerjanya dalam memberikan rekomendasi yang relevan. Pengujian melibatkan 30 responden yang terdiri dari mahasiswa jurusan Teknik Informatika yang sedang dalam proses penulisan skripsi. Para responden diminta untuk memilih topik skripsi berdasarkan rekomendasi yang diberikan oleh sistem. Hasil pengujian diukur dengan metrik kinerja, termasuk presisi, recall, dan F1-score.

Dari 30 rekomendasi yang diberikan, 24 rekomendasi dianggap relevan oleh responden, menghasilkan presisi sistem sebesar 80%. Ini menunjukkan bahwa sebagian besar rekomendasi yang dihasilkan sesuai dengan minat mahasiswa. Namun, dari 40 topik skripsi yang dianggap relevan, sistem berhasil merekomendasikan 24 topik, sehingga nilai recall sistem adalah 60%. Ini menunjukkan bahwa meskipun sistem telah menunjukkan kinerja yang baik, masih ada ruang untuk perbaikan dalam mengidentifikasi topik yang relevan. Nilai F1-score yang diperoleh sebesar 0.69 menunjukkan bahwa sistem sudah cukup baik dalam memberikan rekomendasi, namun masih perlu pengembangan lebih lanjut untuk meningkatkan akurasi dan relevansi rekomendasi.

Analisis hasil menunjukkan bahwa kualitas data yang digunakan sangat mempengaruhi kinerja sistem. Keteraturan dan konsistensi dalam judul dan abstrak skripsi yang ada menjadi faktor penting dalam menentukan efektivitas sistem rekomendasi. Selain itu, penggunaan algoritma *Cosine similarity* dan Jaccard Index terbukti efektif dalam menghitung kesamaan antar dokumen, namun perbandingan dengan metode lain, seperti *collaborative filtering*, dapat memberikan wawasan tambahan untuk meningkatkan akurasi rekomendasi.

Umpan balik dari responden menunjukkan bahwa sebagian besar mahasiswa merasa sistem rekomendasi membantu mereka dalam memilih topik skripsi. Meskipun demikian, beberapa mahasiswa menginginkan variasi lebih dalam rekomendasi yang diberikan, menunjukkan bahwa sistem perlu disempurnakan dengan mempertimbangkan preferensi pengguna yang lebih beragam. Integrasi sistem rekomendasi ke dalam proses akademik di Jurusan Teknik Informatika dapat memberikan manfaat yang signifikan, di mana mahasiswa akan lebih mudah menemukan topik yang sesuai dengan minat mereka, dan jurusan dapat lebih efisien dalam mengelola repositori skripsi serta mengurangi risiko duplikasi judul.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa sistem rekomendasi berbasis *Content-based filtering* yang dikembangkan memiliki potensi untuk memberikan kontribusi positif dalam proses pemilihan topik skripsi di Jurusan Teknik Informatika Universitas Malikussaleh. Meskipun sistem telah menunjukkan kinerja yang baik, masih terdapat beberapa area yang perlu diperbaiki untuk meningkatkan akurasi dan relevansi rekomendasi serta untuk lebih memenuhi kebutuhan mahasiswa.

4. DISKUSI

Hasil penelitian ini menunjukkan bahwa sistem rekomendasi berbasis *Content-based filtering* yang dikembangkan untuk repositori skripsi di Jurusan Teknik Informatika Universitas Malikussaleh telah memberikan kontribusi positif dalam proses pemilihan topik skripsi. Dengan presisi 80% dan recall 60%, sistem ini berhasil menunjukkan kemampuannya dalam memberikan rekomendasi yang relevan bagi mahasiswa. Meskipun demikian, ada beberapa area yang perlu dianalisis lebih dalam untuk memahami kekuatan dan kelemahan sistem yang telah dikembangkan.

Salah satu aspek yang sangat berpengaruh terhadap kinerja sistem adalah kualitas data yang digunakan. Seperti yang telah dicatat, keteraturan dan konsistensi dalam judul dan abstrak skripsi sangat menentukan efektivitas algoritma dalam menghitung kesamaan antar dokumen. Penelitian oleh Kusuma dan Musdholifah (2021) menekankan pentingnya pemeliharaan data yang berkualitas tinggi dalam pengembangan sistem rekomendasi akademik. Hal ini sejalan dengan temuan dalam penelitian ini, di mana data yang tidak terstruktur atau tidak konsisten dapat mengurangi akurasi hasil rekomendasi.

Dari segi metode, penggunaan *Cosine similarity* dan *Jaccard Index* sebagai teknik utama dalam sistem rekomendasi telah terbukti efektif. Hasil penelitian ini sejalan dengan studi Lukas Kurniawan et al. (2023), yang menunjukkan bahwa kombinasi metode berbasis *Content-based filtering* dapat menghasilkan rekomendasi yang akurat. Namun, penelitian sebelumnya juga menunjukkan bahwa penerapan teknik hybrid, yang menggabungkan *Content-based filtering* dengan *collaborative filtering*, dapat meningkatkan performa sistem lebih lanjut. Oleh karena itu, ada potensi untuk mengembangkan sistem ini ke depannya dengan mempertimbangkan pendekatan hybrid guna meningkatkan nilai recall dan variasi rekomendasi yang diberikan.

Umpan balik dari responden menunjukkan bahwa mahasiswa merasa sistem rekomendasi ini membantu mereka dalam memilih topik yang sesuai. Namun, beberapa responden juga mengekspresikan keinginan untuk memiliki lebih banyak variasi dalam rekomendasi yang diberikan. Hal ini mengindikasikan bahwa sistem perlu disempurnakan untuk mempertimbangkan preferensi pengguna yang lebih luas. Penelitian oleh Liga Mayola et al. (2024) menunjukkan bahwa sistem rekomendasi yang responsif terhadap preferensi pengguna dapat menghasilkan hasil yang lebih memuaskan. Oleh karena itu, penerapan metode pembelajaran mesin yang lebih canggih dapat dipertimbangkan untuk lebih memahami pola minat dan preferensi pengguna.

Dalam konteks penerapan akademik, integrasi sistem rekomendasi ini diharapkan dapat meningkatkan efisiensi proses pemilihan topik skripsi di Jurusan Teknik Informatika. Penelitian sebelumnya, seperti yang dilakukan oleh Alifian et al. (2020), mencatat bahwa penerapan teknologi informasi dalam pengelolaan skripsi dapat mempercepat proses administrasi dan meminimalkan terjadinya duplikasi judul. Oleh karena itu, dengan sistem ini, tidak hanya mahasiswa yang diuntungkan, tetapi juga pihak jurusan dalam pengelolaan repositori skripsi.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa sistem rekomendasi berbasis *Content-based filtering* memiliki potensi besar untuk membantu mahasiswa dalam proses pemilihan topik skripsi. Meskipun sistem sudah menunjukkan kinerja yang baik, pengembangan lebih lanjut tetap diperlukan untuk meningkatkan akurasi dan relevansi rekomendasi. Dengan demikian, penelitian ini memberikan kontribusi yang berarti dalam bidang teknologi informasi di lingkungan pendidikan, serta membuka peluang untuk penelitian lanjutan di masa depan.

5. SARAN

Berdasarkan hasil penelitian, disarankan untuk meningkatkan kualitas dan konsistensi data skripsi dalam repositori untuk meningkatkan akurasi sistem rekomendasi. Penggunaan metode hibrid, seperti menggabungkan *content-based filtering* dengan *collaborative filtering*, dapat dipertimbangkan guna memperbaiki variasi rekomendasi dan meningkatkan *recall*. Selain itu, perlu dilakukan pengembangan lebih lanjut untuk mempertimbangkan preferensi pengguna yang lebih luas dan personalisasi dalam rekomendasi.

6. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem rekomendasi berbasis *content-based filtering* untuk repositori skripsi mahasiswa Teknik Informatika Universitas Malikussaleh. Sistem ini menunjukkan kemampuan yang baik dalam memberikan rekomendasi topik skripsi yang relevan dengan presisi sebesar 80% dan *recall* 60%. Meskipun demikian, masih terdapat peluang untuk meningkatkan akurasi, relevansi, dan variasi rekomendasi agar lebih sesuai dengan kebutuhan pengguna. Implementasi sistem ini diharapkan dapat mempermudah mahasiswa dalam memilih topik skripsi serta membantu pihak jurusan dalam mengelola repositori skripsi dengan lebih efisien

7. DAFTAR PUSTAKA

- Hasanah, U., & Mutiara, D. A. (2020). Perbandingan Metode Cosine Similarity dan Jaccard Similarity untuk Penilaian Otomatis Jawaban Pendek. *Sensitif* 2019, 1255–1263.
- Indriani, A., Muhammad, M., Suprianto, S., & Hadriansa, H. (2018). Implementasi *Jaccard Index* Dan N-Gram Pada Rekayasa Aplikasi Koreksi Kata Berbahasa Indonesia. *Sebatik*, 22(2), 95–101. <https://doi.org/10.46984/sebatik.v22i2.314>
- Kurniawan, L., Filiana, A., Virginia, G., & Hartono, B. S. (2023). Sistem Rekomendasi Topik Skripsi Program Studi Informatika. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 9(2), 307. <https://doi.org/10.26418/jp.v9i2.65425>
- Kusuma, H. S., & Musdholifah, A. (2021). Recommendation System for Thesis Topics Using *Content-based filtering*. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(1), 65. <https://doi.org/10.22146/ijccs.62716>
- Leander, J., & Wicaksana, A. (2024). Optimizing a Personalized Movie Recommendation System with Support Vector Machine and *Content-based filtering*. *Journal of System and Management Sciences*, 14(1), 490–501. <https://doi.org/10.33168/JSMS.2024.0128>
- Li, C., Ishak, I., Ibrahim, H., Zolkepli, M., Sidi, F., & Li, C. (2023). Deep Learning-Based Recommendation System: Systematic Review and Classification. *IEEE Access*, 11(September), 113790–113835. <https://doi.org/10.1109/ACCESS.2023.3323353>
- Lumbansiantar, S., Dwiasnati, S., & Fatonah, N. S. (2023). Penerapan Metode Cosine Similarity Dalam Mendeteksi Plagiarisme Pada Jurnal. *Format : Jurnal Ilmiah Teknik Informatika*, 12(2), 142. <https://doi.org/10.22441/format.2023.v12.i2.007>
- Mayola, L., Hafizh, M., & Marse Putra, D. (2024). Algoritma Jaccard Similarity untuk Deteksi Kemiripan Judul Disertasi dengan Pendekatan Variasi Stop Word Removal. *Jurnal Media Informatika Budidarma*, 8(1), 477–487. <https://doi.org/10.30865/mib.v8i1.7109>
- Mayola, L., Hafizh, M., & Putra, D. M. (2024). Perancangan Aplikasi Similarity Deteksi Kemiripan Judul Disertasi Berbasis Web. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 6(2), 452–257. <https://doi.org/10.47233/jteksis.v6i2.1164>
- Melnychuk, T., Galke, L., Seidlmayer, E., Bröring, S., Förstner, K. U., Tochtermann, K., & Schultz, C. (2024). *Development of Similarity Measures From Graph-Structured Bibliographic Metadata : An Application to Identify Scientific Convergence*. 71, 9171–9187.
- Nurchahya, A., & Supriyanto, S. (2020). Content-based recommender system architecture for similar e-commerce products. *Jurnal Informatika*, 14(3), 90. <https://doi.org/10.26555/jifo.v14i3.a18511>
- Prasetyo, D. A. B. (2022). Implementasi Information Schema Database Pada PostgreSQL untuk Pembuatan Tabel Informasi dengan Menggunakan Python Di PT XYZ. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(3), 1961–1972. <https://doi.org/10.35957/jatisi.v9i3.2221>
- Putra Bayu Pratama, N., Mustaqiem, & Minarni. (2021). Pengembangan Sistem Informasi Pengelolaan Judul Skripsi dan Tugas Akhir dengan Fitur Deteksi Kemiripan Menggunakan Algoritma *Winnowing*. *Terapan Informatika Nusantara*, 2(5), 271–278.
- Ramadhan, R. R. (universitas S. N. I. (2024). *Sistem rekomendasi pemesanan menu di kedai our.dish berbasis web dengan metode content based filtering*.