

KOMBINASI *CONTENT-BASED FILTERING* DENGAN *TF-IDF* DAN *COSINE SIMILARITY* UNTUK MEREKOMENDASIKAN TEMPAT WISATA DI SULAWESI UTARA

SKRIPSI

Diajukan untuk memenuhi salah satu
syarat memperoleh gelar Sarjana Komputer
Program Studi Informatika



Disusun Oleh:

Kezia Megumi Manabung

NIM: 215314196

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS SANATA DHARMA
YOGYAKARTA
2025**

**COMBINING CONTENT-BASED FILTERING WITH TF-IDF AND COSINE
SIMILARITY TO RECOMMEND TOURIST ATTRACTIONS IN NORTH
SULAWESI**

SKRIPSI

Submitted to fulfill one of the requirements
for obtaining a Bachelor of Computer Science degree
Study Program: Informatics



Created By:

Kezia Megumi Manabung

NIM: 215314196

FACULTY OF SCIENCE AND TECHNOLOGY

SANATA DHARMA UNIVERSITY

YOGYAKARTA

2025

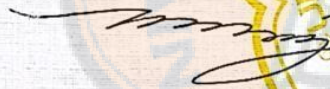
HALAMAN PERSETUJUAN PEMBIMBING

SKRIPSI

KOMBINASI *CONTENT-BASED FILTERING* DENGAN *TF-IDF* DAN *COSINE SIMILARITY* UNTUK MEREKOMENDASIKAN TEMPAT WISATA DI SULAWESI UTARA

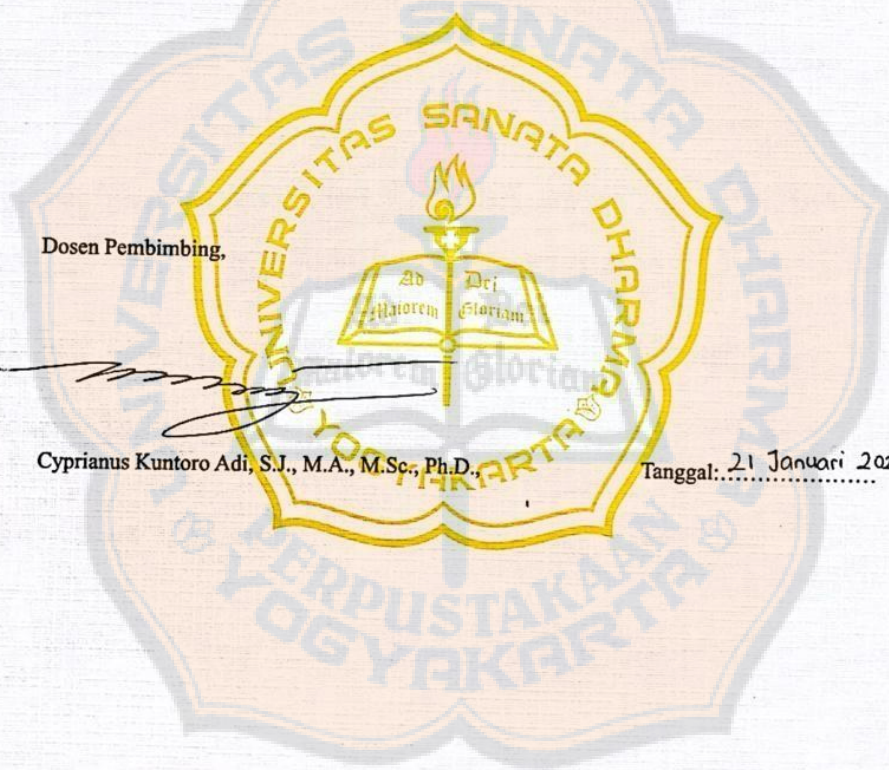
Disusun Oleh:
Kezia Megumi Manabung
NIM: 215314196

Dosen Pembimbing,



Cyprianus Kuntoro Adi, S.J., M.A., M.Sc., Ph.D.

Tanggal: 21 Januari 2025



HALAMAN PENGESAHAN

SKRIPSI

KOMBINASI *CONTENT-BASED FILTERING* DENGAN *TF-IDF* DAN *COSINE SIMILARITY* UNTUK MEREKOMENDASIKAN TEMPAT WISATA DI SULAWESI

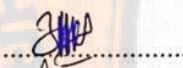
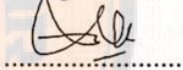
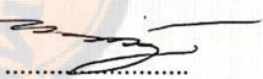
UTARA

Disusun Oleh:

Kezia Megumi Manabung

NIM: 215314196

SUSUNAN DEWAN PENGUJI

JABATAN	NAMA LENGKAP	TANDA TANGAN
Ketua	: Dr. Ir. Iwan Binanto	
Sekretaris	: Agnes Maria Polina, S.kom., M.Sc.	
Anggota	: Cyprianus Kuntoro Adi, S.J., M.A., M.Sc., Ph.D.,	

Yogyakarta, 21 Januari 2025

Fakultas Sains dan Teknologi

Universitas Sanata Dharma

Dekan,



Ir. Drs. Haris Sriwindono, M.Kom, Ph.D.

HALAMAN PERNYATAAN KEASLIAN KARYA

Saya menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis ini tidak memuat karya atau bagian karya orang lain, kecuali yang telah disebutkan dalam kutipan dan daftar pustaka dengan mengikuti ketentuan sebagaimana layaknya karya ilmiah.

Apabila di kemudian hari ditemukan indikasi plagiarisme dalam naskah ini, saya bersedia menanggung segala sanksi sesuai peraturan perundang-undangan yang berlaku.

Yogyakarta, 21 Januari 2025

Penulis,



Kezia Megumi Manabung

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA
ILMIAH UNTUK KEPERLUAN AKADEMIS**

Yang bertanda tangan di bawah ini, saya mahasiswa Universitas Sanata Dharma:

Nama : Kezia Megumi Manabung

NIM : 215314196

Demi perkembangan ilmu pengetahuan, saya memberikan kepada Perpustakaan Universitas Sanata Dharma karya ilmiah saya yang berjudul:

**“KOMBINASI *CONTENT-BASED FILTERING* DENGAN *TF-IDF* DAN *COSINE SIMILARITY* UNTUK MEREKOMENDASIKAN TEMPAT WISATA DI
SULAWESI UTARA”**

beserta perangkat yang diperlukan (bila ada). Dengan demikian saya memberikan hak kepada Perpustakaan Universitas Sanata Dharma baik untuk menyimpan, mengalihkan dalam bentuk media lain, mengolah dalam bentuk pangkalan data, mendistribusikan secara terbatas, dan mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya atau memberikan royalti kepada saya selama tetap mencantumkan nama saya sebagai penulis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Yogyakarta
Pada tanggal: 21 Januari 2025
Yang menyatakan,



Kezia Megumi Manabung

MOTTO

“Carpe Diem, menikmati waktu yang ada sekarang tanpa terlalu khawatir tentang masa depan, serta menghargai setiap momen dalam hidup”

-Kezia Megumi Manabung



KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala penyertaan dan bimbingan-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul **"KOMBINASI CONTENT-BASED FILTERING DENGAN TF-IDF DAN COSINE SIMILARITY UNTUK MEREKOMENDASIKAN TEMPAT WISATA DI SULAWESI UTARA"** Penyusunan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Universitas Sanata Dharma.

Selama proses penyusunan skripsi ini, penulis menyadari bahwa capaian ini tidak lepas dari dukungan, bantuan, dan doa dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis ingin menyampaikan apresiasi kepada:

1. Tuhan Yesus Kristus, yang senantiasa memberikan kekuatan, hikmat, dan kesehatan sehingga penulis dapat menyelesaikan skripsi ini.
2. Kedua orang tua penulis, Mama dan Daddy, serta seluruh anggota keluarga yang telah memberikan doa, cinta, dan dukungan moral maupun material selama proses perkuliahan hingga penyusunan skripsi ini.
3. Romo Cyprianus Kuntoro Adi, S.J., M.A., M.Sc., Ph.D., selaku dosen pembimbing yang dengan penuh kesabaran dan ketelitian telah membimbing penulis selama proses penelitian dan penulisan skripsi ini.
4. Seluruh dosen Program Studi Informatika, yang telah memberikan ilmu, bimbingan, dan motivasi kepada penulis selama masa perkuliahan hingga penyelesaian skripsi ini.
5. Semua pihak lain, baik yang membantu secara langsung maupun tidak langsung, yang telah memberikan kontribusi dalam bentuk apa pun selama proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih belum sempurna dan terbuka terhadap kritik serta saran yang membangun. Penulis berharap karya ini bermanfaat bagi pembaca, terutama dalam pengembangan sistem rekomendasi berbasis teknologi informasi.

Yogyakarta, 06 Januari 2025
Penulis,



Kezia Megumi Manabung

ABSTRAK

Sulawesi Utara merupakan salah satu provinsi di Indonesia yang memiliki potensi pariwisata yang kaya akan keindahan alam dan budaya. Namun, banyaknya pilihan destinasi wisata sering kali menyulitkan wisatawan dalam menentukan tujuan yang sesuai dengan preferensi mereka. Untuk mengatasi permasalahan tersebut, penelitian ini bertujuan untuk mengembangkan sistem rekomendasi tempat wisata di Sulawesi Utara dengan menggunakan pendekatan *Content-Based Filtering* (CBF).

Metode CBF yang digunakan dalam penelitian ini memanfaatkan pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk menganalisis kata-kata dalam deskripsi objek wisata yang telah diproses sebelumnya. Selain itu, sistem ini menghitung tingkat kesamaan antar deskripsi objek wisata dengan preferensi pengguna menggunakan metode cosine similarity. Objek wisata yang memiliki tingkat kesamaan tertinggi dengan preferensi pengguna akan direkomendasikan.

Penelitian ini menggunakan dataset yang berisi nama objek wisata, lokasi, dan deskripsi sebagai basis data sistem. Hasil penelitian menunjukkan bahwa metode TF-IDF dan *cosine similarity* dapat meningkatkan akurasi sistem rekomendasi dengan menyajikan tempat wisata yang relevan sesuai preferensi pengguna. Sistem ini diharapkan mampu membantu wisatawan dalam memilih destinasi wisata dengan lebih mudah dan meningkatkan pengalaman mereka selama berwisata di Sulawesi Utara.

Untuk mengevaluasi kinerja sistem, dilakukan tiga uji berbeda: uji *precision*, uji *recall*, dan uji validitas. Uji *precision* dan *recall* dilakukan menggunakan kuisisioner dengan 8 responden, yang menghasilkan nilai *precision* sebesar 80% dan *recall* sebesar 36,758%. Selain itu, uji validitas dilakukan dengan membandingkan hasil perhitungan manual dan perhitungan sistem, yang menunjukkan hasil akurasi 100%. Sistem ini diharapkan mampu membantu wisatawan dalam memilih destinasi wisata dengan lebih mudah dan meningkatkan pengalaman mereka selama berwisata di Sulawesi Utara.

Kata Kunci: *Content-Based Filtering*, TF-IDF, *cosine similarity*, sistem rekomendasi, Sulawesi Utara, tempat wisata, destinasi

ABSTRACT

North Sulawesi is one of Indonesia's provinces rich in natural beauty and cultural heritage. However, the abundance of tourist destination options often makes it challenging for travelers to choose destinations that align with their preferences. To address this issue, this study aims to develop a tourist attraction recommendation system in North Sulawesi using the Content-Based Filtering (CBF) approach.

The CBF method employed in this study utilizes TF-IDF (Term Frequency-Inverse Document Frequency) weighting to analyze the processed textual descriptions of tourist attractions. Furthermore, the system calculates the similarity between tourist attraction descriptions and user preferences using the cosine similarity method. Tourist attractions with the highest similarity scores to the user preferences are recommended.

This research utilizes a dataset containing tourist attraction names, locations, and descriptions as the system's database. The findings demonstrate that the TF-IDF and cosine similarity methods significantly improve the recommendation system's accuracy by providing relevant tourist destinations based on user preferences. This system is expected to assist tourists in selecting destinations more efficiently and enhance their overall travel experiences in North Sulawesi.

To evaluate the system's performance, three different tests were conducted: precision test, recall test, and validity test. The precision and recall tests were performed using a questionnaire with 8 respondents, resulting in a precision value of 80% and a recall value of 36.758%. Additionally, the validity test was carried out by comparing the results of manual calculations with those of the system, yielding an accuracy of 100%. This system is expected to assist tourists in selecting tourist destinations more easily and enhance their experience while traveling in North Sulawesi.

Keywords: Content-Based Filtering, TF-IDF, cosine similarity, recommendation system, North Sulawesi, tourist attractions, destinations.

DAFTAR PUSTAKA

- (n.d.). Retrieved from Boltim PikiranRakyat: <https://boltim.pikiran-rakyat.com/>
- (n.d.). Retrieved from TribunManadotravel.com: <https://tribunmanadotravel.tribunnews.com/>
- 54 *Tempat Wisata di Sulawesi Utara, Poppuler hingga tersembunyi.* (n.d.). Retrieved from Senang Rekreasi: https://www.senangrekreasi.com/tempat-wisata-sulawesi-utara/#google_vignette
- Aminah, S. (n.d.). Perbandingan Keakuratan Sistem Rekomendasi Produk Berbasis Content-Based Filtering Dan Collaborative Filtering Pada E-Commerce Shopee Menggunakan Matrik Precision,. *Online Repository of Universitas NU Kalimantan Selatan*, 11-13.
- BAB III *Landasan Teori.* (n.d.). Retrieved from https://e-journal.uajy.ac.id/8852/4/3TF06448.pdf?utm_source
- Content Based Filtering dalam Algoritma Data Science.* (2023, November 28). Retrieved from DQLab: <https://dqlab.id/content-based-filtering-dalam-algoritma-data-science>
- Content Based Filtering dalam Algoritma Data Science.* (2023, November 28). Retrieved from DQLab: <https://dqlab.id/content-based-filtering-dalam-algoritma-data-science>
- Dewi, A. C. (2021, Februari 6). *Klasifikasi menggunakan Algoritma Decision Tree.* Retrieved from Medium: <https://agneschintiadewi.medium.com/klasifikasi-menggunakan-algoritma-decision-tree-446d500ba73c>
- Fajriansyah, M., Adikara, P. P., & Widodo, A. W. (2021). Sistem Rekomendasi Film Menggunakan Content Based Filtering . *JURNAL PENGEMBANGAN TEKNOLOGI INFORMASI DAN ILMU KOMPUTER.*
- Gradiano, R. A. (2022, September 27). *Pengertian Pariwisata Menurut Para Ahli, Pahami Tujuan dan Manfaatnya.* Retrieved from bola.com: <https://www.bola.com/ragam/read/5081289/pengertian-pariwisata-menurut-para-ahli-pahami-tujuan-dan-manfaatnya>
- Huda, A. A., Fajarudin, R., & Hadinegoro, A. (2022). Sistem Rekomendasi Content-based Filtering Menggunakan TF-IDF. *Building of Informatics, Technology and Science (BITS).*

- Ibnu Surya Wibowo, A. W. (2024). Keyword Extraction Judul Berita Online Di Indonesia Menggunakan Metode TF-IDF. *Jurnal Teknik Informatika dan Sistem Informasi*. Retrieved from <https://jurnal.mdp.ac.id/index.php/jatisi/article/view/6718/1758>
- Keindahan Desa Wisata Pulisan di Minahasa Utara*. (2022, Mei 17). Retrieved from Kementerian Pariwisata dan Ekonomi Kreatif: <https://kemenparekraf.go.id/ragam-pariwisata/keindahan-desa-wisata-pulisan-di-minahasa-utara>
- Kelong Café & Resto*. (n.d.). Retrieved from LOSNITO: <https://www.losnito.com/indahnya-taman-kelong-cafe-resto/>
- Muryani. (2020). ANALISIS SEKTOR PARIWISATA DAN DAMPAK PENGELUARAN WISATAWAN TERHADAP PEREKONOMIAN PROVINSI SULAWESI UTARA. *Jurnal Ekonomi dan Keuangan*, 123.
- Nurdiana, O., Jumadi, & Nursantika, D. (2016). Perbandingan Metode Cosine Similarity Dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia. *Jurnal Online Informatika*.
- Oktafiani, R., & Rianto, R. (2023). Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree untuk Sistem Rekomendasi Tempat Wisata. *JURNAL NASIONAL TEKNOLOGI & SISTEM INFORMASI*.
- Parwita, W. G. (2019, Februari 1). Pengujian Akurasi Sistem Rekomendasi Berbasis Content-Based Filtering. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*. Retrieved from <https://repository.unukase.ac.id/id/eprint/268/1/PERBANDINGAN%20KEAKURATAN%20SISTEM%20REKOMENDASI%20PRODUK%20BERBASIS%20CONTENT-BASED%20FILTERING%20DAN%20COLLABORATIVE%20FILTERING%20PADA%20E-COMMERCE%20SHOPEE%20MENGUNAKAN%20MATRIK%20PRECISION%20C%20RECALL>
- Pramesti, D. A., & Santiyasa, I. (2022). Penerapan Metode Content-Based Filtering dalam Sistem Rekomendasi Video Game. *Jurnal Harian Regional*.
- Rekomendasi Objek Wisata Sulawesi Utara*. (n.d.). Retrieved from TribunManado.co.id.: <https://manado.tribunnews.com/>
- Salim, E., Pragantha, J., & Lauro, M. D. (n.d.). Perancangan Sistem Rekomendasi Film menggunakan metode Content-Based Filtering. *Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara*.

Supiyanto, & Sriyono. (2023). Metode Cosine Similarity Untuk Mendeteksi Kemiripan Pada Dokumen Teks. *SAINS Jurnal MIPA dan Pengajarannya*.

Tempat Wisata di Sulawesi Utara. (n.d.). Retrieved from Tripadvisor: <https://www.tripadvisor.co.id/>

Wisata Sulawesi Utara. (n.d.). Retrieved from KOMPAS.com: <https://www.kompas.com/>

Wisata Terbaik di Sulawesi Utara. (n.d.). Retrieved from atourin: <https://>

