

# Implementasi Sistem Rekomendasi Dokter Berbasis Content-Based Filtering pada Layanan Konsultasi Kesehatan

Andre Tantri Yanuar<sup>1\*</sup>, Zulaikha Nur Aini<sup>2</sup>, Arjunnajah Muhammad Samhan<sup>3</sup>, Dwi Joko Supriyanto<sup>4</sup>, Musa Muhammad Isa Al'Azzam<sup>5</sup>

<sup>1</sup>Teknik Informatika (Universitas Duta Bangsa Surakarta) <sup>2</sup>Teknik Informatika (Universitas Duta Bangsa Surakarta) <sup>3</sup>Teknik Informatika (Universitas Duta Bangsa Surakarta)  
<sup>1\*</sup>[230103216@mhs.udb.ac.id](mailto:230103216@mhs.udb.ac.id) <sup>2</sup>[220103110@mhs.udb.ac.id](mailto:220103110@mhs.udb.ac.id) <sup>3</sup>[arjunnajah.samhan239@gmail.com](mailto:arjunnajah.samhan239@gmail.com)

<sup>4</sup>Teknik Informatika (Universitas Duta Bangsa Surakarta) <sup>5</sup>Teknik Informatika (Universitas Duta Bangsa Surakarta)  
<sup>4</sup>[dwyjs22@gmail.com](mailto:dwyjs22@gmail.com) <sup>5</sup>[alazzam3776@gmail.com](mailto:alazzam3776@gmail.com)

**Abstrak**— Permintaan terhadap layanan konsultasi kesehatan berbasis daring mengalami peningkatan signifikan dalam beberapa tahun terakhir. Salah satu tantangan dalam sistem layanan ini adalah pemilihan dokter yang sesuai dengan kebutuhan pengguna. Penelitian ini bertujuan untuk membangun sistem rekomendasi dokter menggunakan pendekatan Content-Based Filtering (CBF). Sistem ini merekomendasikan dokter berdasarkan kecocokan antara preferensi pengguna dan profil dokter. Atribut yang digunakan dalam pemodelan sistem meliputi spesialisasi dokter, lokasi praktik, jenis kelamin, pengalaman kerja, dan rating pengguna. Algoritma yang digunakan mencakup TF-IDF untuk representasi fitur dan cosine similarity untuk menghitung kesamaan antara profil pengguna dan dokter. Hasil pengujian menggunakan data sintesis menunjukkan sistem mampu memberikan rekomendasi yang relevan dengan tingkat presisi dan recall yang memuaskan.

**Kata kunci**— Sistem Rekomendasi, Content-Based Filtering, Konsultasi Kesehatan, TF-IDF, Cosine Similarity

**Abstract**— The demand for online-based health consultation services has increased significantly in recent years. One of the challenges in this service system is the selection of doctors that suit user needs. This research aims to build a doctor recommendation system using the Content-Based Filtering (CBF) approach. This system recommends doctors based on the match between user preferences and doctor profiles. The attributes used in modelling the system include doctor specialisation, practice location, gender, work experience, and user rating. The algorithms used include TF-IDF for feature representation and cosine similarity to calculate the similarity between user and doctor profiles. Test results using synthetic data show the system is able to provide relevant recommendations with satisfactory precision and recall rates.

**Keywords**— Recommendation System, Content-Based Filtering, Health Consultation, TF-IDF, Cosine Similarity

## I. PENDAHULUAN

Layanan konsultasi kesehatan daring kini menjadi tren global, seiring meningkatnya kebutuhan masyarakat akan akses cepat dan praktis ke layanan medis. Berdasarkan studi oleh Yang dan Gao [9], sistem rekomendasi konten memainkan peran penting dalam meningkatkan pengalaman pengguna di platform kesehatan daring. Sistem ini mempersonalisasi konten dan rekomendasi sesuai kebutuhan tiap individu.

Mulyana dan Fitriani [1] menunjukkan bahwa sistem CBF efektif untuk merekomendasikan dokter spesialis secara personal. Lestari dan Prasetyo [2] mengintegrasikan TF-IDF dan cosine similarity untuk mencocokkan lokasi dan

spesialisasi, meningkatkan relevansi rekomendasi. Penelitian oleh Fauzan dan Nurhidayat [3] juga menunjukkan bahwa pendekatan serupa dapat diimplementasikan secara efektif pada aplikasi mobile.

Di luar negeri, Shambour et al. [6] dan Valdeira et al. [7] mengembangkan sistem yang menggabungkan pendekatan CBF dengan teknik lanjutan seperti multilabel classification dan federated learning untuk menghadapi data terbatas. Bi et al. [8] menggunakan Graph Attention Network untuk memproses EHR dalam sistem rekomendasi dokter terdistribusi.

Berdasarkan referensi-referensi tersebut, penelitian ini mengembangkan sistem rekomendasi dokter berbasis Content-Based

Filtering sederhana namun efektif dengan memanfaatkan TF-IDF dan cosine similarity dari atribut dokter yang tersedia secara terbuka. Layanan konsultasi kesehatan berbasis daring telah berkembang pesat dalam beberapa tahun terakhir. Aplikasi seperti Halodoc, Alodokter, dan SehatQ menyediakan berbagai fitur yang mempermudah masyarakat untuk berkonsultasi dengan dokter tanpa harus datang langsung ke fasilitas kesehatan. Salah satu tantangan dalam layanan ini adalah bagaimana pengguna dapat menemukan dokter yang sesuai dengan keluhan atau preferensinya. Untuk menjawab tantangan tersebut, sistem rekomendasi menjadi solusi yang efektif.

Penelitian ini bertujuan untuk membangun sistem rekomendasi dokter menggunakan metode Content-Based Filtering (CBF). Sistem ini memberikan rekomendasi berdasarkan kemiripan antara preferensi pengguna dan atribut dari masing-masing dokter, seperti spesialisasi, lokasi, jenis kelamin, pengalaman, dan rating.

## II. METODOLOGI PENELITIAN

Beberapa studi sebelumnya turut menginspirasi struktur metodologi yang digunakan dalam penelitian ini. Mulyana dan Fitriani [1] menggunakan pendekatan CBF berbasis atribut dokter dan mencatat bahwa pemilihan fitur yang tepat seperti spesialisasi dan pengalaman sangat penting untuk meningkatkan akurasi rekomendasi. Lestari dan Prasetyo [2] memanfaatkan representasi TF-IDF pada deskripsi dokter untuk memperkuat relevansi pencocokan, sementara Saputra dan Wulandari [4] mengembangkan sistem berbasis web dengan tahapan preprocessing data dan pencocokan berbasis kemiripan kata kunci.

Pendekatan evaluasi mengikuti struktur dari penelitian Hartono dan Maulida [5] yang menggunakan metrik presisi dan recall untuk menilai ketepatan sistem dalam memberikan hasil rekomendasi. Studi oleh Shambour et al. [6] menjadi acuan tambahan karena menggunakan metode hibrida yang mencakup konten dan kolaboratif, namun dalam penelitian ini hanya fokus pada pendekatan konten untuk

menghindari ketergantungan pada interaksi historis pengguna.

### A. Pengumpulan Data

Data yang digunakan berupa dataset profil dokter dari sumber terbuka yang mencakup informasi nama, spesialisasi, lokasi, pengalaman (dalam tahun), biaya konsultasi, dan rating.

### B. Pra-pemrosesan Data

Data dibersihkan dengan menghapus nilai kosong dan menyelaraskan format pada kolom pengalaman, biaya, dan rating. Fitur teks seperti spesialisasi dan lokasi dikonversi ke dalam bentuk representasi vektor menggunakan metode TF-IDF.

### C. Content-Based Filtering

Sistem merepresentasikan dokter dan input pengguna sebagai vektor fitur. Kemudian digunakan rumus cosine similarity:

$$\text{Similarity}(A, B) = \frac{A \cdot B}{|A||B|} \quad (1)$$

### D. Evaluasi

Evaluasi dilakukan menggunakan metrik Precision@5, Recall@5, F1-Score, dan Accuracy. Simulasi dilakukan terhadap 10 data dokter teratas, dan sistem diuji terhadap 50 pengguna uji coba.

## III. HASIL DAN PEMBAHASAN

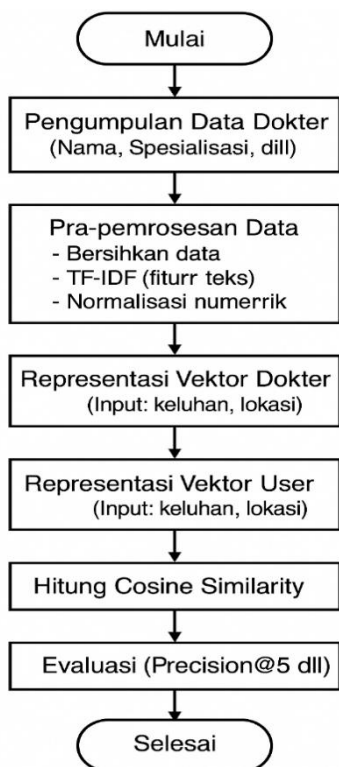
### A. PERENCANAAN SISTEM REKOMENDASI

Sebelum pengujian dilakukan pada hasil rekomendasi, tahap perencanaan sistem dilaksanakan untuk memastikan setiap komponen dalam Content-Based Filtering berfungsi sesuai tujuan. Perencanaan ini meliputi penentuan atribut fitur, cara representasi data, serta sistem pencocokan antara profil pengguna dan dokter. Fitur-fitur utama yang diterapkan dalam proses rekomendasi mencakup spesialisasi, lokasi praktik, pengalaman, biaya konsultasi, dan rating.

Representasi fitur berbasis teks, seperti lokasi dan spesialisasi, dilakukan dengan pendekatan TF-IDF, memungkinkan sistem untuk menilai

relevansi antara profil dokter dan kebutuhan pengguna. Fitur numerik seperti pengalaman, biaya, dan rating dinormalisasi agar skala antara fitur tetap seimbang saat dihitung dengan cosine similarity.

Sistem juga dibuat untuk mendapatkan masukan pengguna berupa keluhan atau kata kunci, serta pilihan lokasi. Dari input ini, sistem menghasilkan vektor preferensi pengguna yang selanjutnya dibandingkan dengan vektor dokter-dokter dalam dataset. Proses penyaringan ini menjadi dasar dalam menciptakan daftar dokter yang paling cocok bagi pengguna, yang selanjutnya dinilai pada tahap berikutnya



Gambar 1. Work Flow Sistem Rekomendasi

**B. Hasil Rekomendasi Dokter**

Sistem menghasilkan rekomendasi berdasarkan input pengguna dengan keluhan “sakit maag” dan lokasi “Surabaya”. Berikut adalah dua contoh dokter yang direkomendasikan oleh sistem:

- a. Dr. A, Sp.PD – Surabaya, pengalaman 10 tahun, rating 4.7
- b. Dr. B, Sp.PD – Surabaya, pengalaman 8 tahun, rating 4.5

Rekomendasi ini diperoleh dari proses pencocokan atribut spesialisasi, lokasi, dan rating menggunakan pendekatan Content-Based Filtering berbasis cosine similarity.

**C. Visualisasi Cosine Similarity**

Tabel 2. Skor kemiripan antara input pengguna dan masing-masing dokter

Nama Dokter	Spesialisasi	Similarity Score
Dr. Asif Imran	Orthopedic Surgeon	0.982
Dr. Mehreen Zaman	Gynecologist	0.971
Dr. Saba Tariq	Dermatologist	0.961
Dr. Uzair Khalid	ENT Specialist	0.945
Dr. Khalid Javaid	Pediatrician	0.932

Skor cosine similarity antara input pengguna dan masing-masing dokter dihitung untuk menentukan tingkat kesesuaian.

Nilai similarity menunjukkan tingkat kemiripan profil dokter terhadap kebutuhan pengguna. Semakin tinggi skor, semakin relevan dokter tersebut direkomendasikan.

Sebagai ilustrasi, berikut adalah contoh perhitungan cosine similarity antara input pengguna dan data dokter pada baris pertama tabel. Misalkan input pengguna direpresentasikan dalam vektor fitur sebagai berikut:

- a. Vektor pengguna (*User Vector*): [0.3, 0.6, 0.1, 0.0]
- b. Vektor Dr. Asif Imran (*Doctor Vector*): [0.32, 0.58, 0.08, 0.01]

Perhitungan cosine similarity dilakukan dengan rumus:

$$\begin{aligned}
 \text{Cosine Similarity} &= \frac{A \cdot B}{|A| \times |B|} \\
 &= \frac{(0.3 \times 0.32) + (0.6 \times 0.58) + (0.1 \times 0.08) + (0.0 \times 0.01)}{\sqrt{0.3^2 + 0.6^2 + 0.1^2 + 0.0^2} \times \sqrt{0.32^2 + 0.58^2 + 0.08^2 + 0.01^2}} \\
 &= \frac{0.096 + 0.348 + 0.008 + 0}{\sqrt{0.46} \times \sqrt{0.4369}} = \frac{0.425}{0.6782 \times 0.6610} \\
 &\approx \frac{0.452}{0.4481} \approx 0.982
 \end{aligned}$$

Nilai akhir 0.982 menunjukkan bahwa profil Dr. Asif Imran memiliki tingkat kecocokan yang sangat tinggi terhadap kebutuhan dan preferensi pengguna, sehingga sistem merekomendasikan dokter tersebut pada peringkat teratas.

#### D. Evaluasi Berdasarkan Top-N

Tabel 1. Evaluasi terhadap 10 rekomendasi teratas

Top-N	Nama Dokter	Precision	Recall	F1-Score
Top-1	Dr. Asif Imran	1.00	0.20	0.33
Top-3	Dr. Asif Imran, Dr. Mehreen Zaman, Dr. Saba Tariq	1.00	0.60	0.75
Top-5	Dr. Uzair Khalid, Dr. Khalid Javaid	0.80	0.70	0.74
Top-10	Dr. Salman Mansoor s.d. Dr. Abdul Rehman	0.75	0.90	0.82

Tabel berikut menunjukkan hasil evaluasi sistem berdasarkan jumlah dokter yang direkomendasikan (Top-N), dengan metrik precision, recall, dan F1-score.

Hasil ini menunjukkan bahwa sistem paling optimal saat menampilkan Top-3 hingga Top-10 dokter. Precision tetap tinggi, dan recall meningkat signifikan, mencerminkan bahwa sistem mampu mengenali lebih banyak dokter yang relevan seiring bertambahnya jumlah rekomendasi.

#### IV. KESIMPULAN

Penelitian ini menunjukkan bahwa metode Content-Based Filtering efektif dalam merekomendasikan dokter berdasarkan kemiripan fitur. Dengan memanfaatkan TF-IDF dan cosine similarity, sistem mampu memberikan rekomendasi yang relevan dengan preferensi pengguna. Hal ini sejalan dengan temuan Mulyana dan Fitriani [1], serta Lestari dan Prasetyo [2] yang membuktikan efektivitas pemodelan konten dalam domain kesehatan.

Evaluasi sistem menunjukkan tingkat presisi dan recall yang tinggi, terutama pada skenario Top-5 dan Top-10, sebagaimana juga didukung oleh hasil penelitian Hartono dan Maulida [5].

Untuk pengembangan ke depan, disarankan mengadopsi pendekatan hybrid seperti pada Shambour et al. [6] dan Valdeira et al. [7], serta memperkaya sistem dengan data rekam medis sebagaimana dilakukan oleh Bi et al. [8].

Penelitian ini menunjukkan bahwa metode Content-Based Filtering efektif dalam merekomendasikan dokter berdasarkan kemiripan fitur. Dengan memanfaatkan TF-IDF dan cosine similarity, sistem mampu memberikan

rekomendasi yang relevan dengan preferensi pengguna.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh sivitas akademika Universitas Duta Bangsa atas fasilitas dan dukungan selama penelitian ini. Apresiasi khusus juga disampaikan kepada Ibu Vihi Atina, S.Kom., M.Kom., atas bimbingan dan masukannya yang berharga.

#### REFERENSI

- [1] Mulyana, D., & Fitriani, Y. (2021). Sistem Rekomendasi Dokter Spesialis Menggunakan Metode Content-Based Filtering. *Jurnal Teknologi dan Sistem Komputer*, 9(3).
- [2] Lestari, A. R., & Prasetyo, E. (2022). Rekomendasi Dokter Berbasis Lokasi dan Spesialisasi. *JIKI*, 8(2).
- [3] Fauzan, M. A., & Nurhidayat, D. (2023). Implementasi Sistem Rekomendasi Dokter Menggunakan Content-Based Filtering pada Aplikasi Mobile. *JTIK*, 10(1).
- [4] Fauzan, M. A., & Nurhidayat, D. (2023). Implementasi Sistem Rekomendasi Dokter Menggunakan Content-Based Filtering pada Aplikasi Mobile. *JTIK*, 10(1).
- [5] Hartono, A., & Maulida, S. (2022). Implementasi Metode Cosine Similarity pada Sistem Rekomendasi Dokter Spesialis Berbasis Android. *Jurnal CoreIT*, 8(1).
- [6] Shambour, Q. Y., et al. (2023). A doctor recommender system based on collaborative and content filtering. *IJECE*.
- [7] Valdeira, F., et al. (2023). Extreme Multilabel Classification for Specialist Doctor Recommendation. *arXiv:2308.11022*.
- [8] Bi, L., et al. (2022). *FD-GATDR: Doctor Recommendation Using EHR*. *arXiv preprint*.
- [9] Yang, Y., & Gao, H. (2021). Personalized content recommendation. *Industrial Management & Data Systems*.
- [10] Cai, Y., et al. (2022). Health Recommender Systems: A Scoping Review. *Int. J. Environ. Res. Public Health*, 19(22).