

PENERAPAN METODE *CONTENT-BASED FILTERING* PADA SISTEM REKOMENDASI *COFFEE SHOP* DI KOTA SURAKARTA BERBASIS WEB

Sahrul Ageng Prasetyo, Pramono, Vihi Atina

Teknik Informatika, Universitas Duta Bangsa Surakarta

Jl. Bhayangkara No.55, Tipes, Kec. Serengan, Kota Surakarta, Jawa Tengah 57154

prasetyoarul0@gmail.com

ABSTRAK

Kota Surakarta merupakan salah satu pusat pendidikan dan budaya yang memiliki pertumbuhan industri kreatif kuliner sangat pesat. Namun, menjamurnya jumlah *coffee shop* saat ini menimbulkan fenomena *information overload* yang menyulitkan konsumen dalam memilih lokasi yang paling sesuai dengan kriteria personal mereka. Penelitian ini bertujuan untuk merancang dan membangun sistem rekomendasi *coffee shop* berbasis web yang mampu memberikan pilihan keputusan secara cepat dan akurat. Pengembangan sistem ini menggunakan model *prototype*. Logika rekomendasi dikembangkan dengan menerapkan metode *Content-Based Filtering* yang mengintegrasikan teknik pembobotan TF-IDF untuk ekstraksi fitur atribut *coffee shop* dan perhitungan *cosine similarity* untuk mengukur tingkat kemiripan karakteristiknya. Hasil pengujian fungsional menunjukkan bahwa sistem web dapat berjalan dengan baik sesuai kebutuhan pengguna. Berdasarkan evaluasi performa algoritma menggunakan *confusion matrix* terhadap 100 data *coffee shop*, sistem berhasil memperoleh rata-rata *precision* sebesar 88,00%, rata-rata *recall* mencapai 88,67%, serta *F1-score* sebesar 87,96%. Hasil tersebut membuktikan bahwa metode yang diimplementasikan memiliki tingkat akurasi yang tinggi dalam menghasilkan rekomendasi yang relevan.

Kata kunci : sistem rekomendasi, *Content-Based Filtering*, *coffee shop*, TF-IDF, *cosine similarity*, web

1. PENDAHULUAN

Perkembangan industri kuliner, khususnya di Kota Surakarta, mengalami peningkatan yang sangat signifikan dalam beberapa tahun terakhir. Banyaknya *coffee shop* yang bermunculan dengan menawarkan berbagai konsep ruang, kelengkapan fasilitas fungsional, dan variasi menu memberikan beragam pilihan bagi masyarakat dalam menentukan tempat yang representatif untuk bersantai atau bekerja. Namun, pesatnya pertumbuhan dan banyaknya pilihan tersebut justru menimbulkan kesulitan baru bagi konsumen dalam menentukan lokasi yang benar-benar sesuai dengan preferensi personal mereka, sehingga sering terjadi fenomena *information overload* yang menghambat proses pengambilan keputusan [1].

Permasalahan penumpukan informasi tersebut menunjukkan perlunya sebuah sistem cerdas yang mampu membantu pengguna dalam menyaring informasi secara personal dan memberikan rekomendasi yang relevan. Sistem rekomendasi merupakan solusi teknologi tepat guna yang dapat diimplementasikan untuk memberikan saran keputusan berdasarkan kecocokan preferensi pengguna. Salah satu metode yang efektif untuk diterapkan adalah *Content-Based Filtering*, yaitu metode pemfilteran yang memberikan rekomendasi item baru berdasarkan tingkat kemiripan atribut dari suatu item dengan karakteristik preferensi yang diinginkan oleh pengguna [2].

Sebagai rencana penyelesaian masalah yang diangkat, metode *Content-Based Filtering* dalam penelitian ini diterapkan dengan memanfaatkan teknik *Term Frequency-Inverse Document Frequency* untuk melakukan pembobotan tingkat kepentingan atribut

coffee shop, serta algoritma *Cosine Similarity* untuk menghitung nilai kedekatan atau tingkat kemiripan antar *coffee shop*. Melalui integrasi pendekatan ini, sistem dibangun agar dapat memberikan rekomendasi yang jauh lebih personal dan objektif, serta mampu berjalan secara optimal tanpa ketergantungan pada ketersediaan data penilaian dari pengguna lain [3].

Melalui penerapan algoritma tersebut, penelitian ini bertujuan untuk merancang dan membangun sebuah sistem rekomendasi *coffee shop* berbasis web di Kota Surakarta dengan menggunakan metode pengembangan perangkat lunak *Prototype*. Dengan diimplementasikannya sistem ini, hasil akhir yang dikembangkan diharapkan mampu bertindak sebagai solusi *aplikatif* yang membantu masyarakat maupun pendatang di Kota Surakarta dalam menemukan *coffee shop* yang sesuai dengan preferensi mereka secara lebih cepat, tepat, dan efisien [4].

2. TINJAUAN PUSTAKA

Sebelum melakukan penelitian ini, penulis terlebih dahulu melakukan studi literatur dengan menelusuri berbagai referensi yang relevan, baik berupa buku maupun jurnal ilmiah nasional yang berkaitan dengan sistem rekomendasi, metode *Content-Based Filtering*, serta teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *cosine similarity*. Setelah dilakukan studi literatur, diperoleh beberapa penelitian terdahulu yang relevan dengan topik sistem rekomendasi. Mengembangkan sistem rekomendasi buku perpustakaan sekolah dengan menerapkan metode *Content-Based Filtering* menggunakan TF-IDF dan *cosine similarity*. Metode tersebut digunakan untuk memberikan rekomendasi

buku berdasarkan kemiripan isi atau kategori buku yang pernah dipilih oleh pengguna. Hasil penelitian menunjukkan bahwa sistem mampu memberikan rekomendasi yang relevan bagi pengguna [5].

Selanjutnya, mengembangkan sistem rekomendasi *coffee shop* di Kota Surakarta dengan tujuan membantu pengguna menemukan *coffee shop* yang sesuai dengan preferensi pengguna [6]. Sistem ini menggunakan pendekatan item-based collaborative filtering yang bekerja dengan menghitung tingkat kemiripan antar item berdasarkan riwayat interaksi atau penilaian pengguna. Hasil evaluasi menunjukkan bahwa sistem memiliki tingkat akurasi yang cukup tinggi berdasarkan pengujian menggunakan confusion matrix, sehingga metode ini dinilai efektif dalam menghasilkan rekomendasi yang relevan.

Penelitian lain berfokus pada permasalahan kesulitan konsumen dalam memilih produk perawatan kulit yang sesuai dengan kondisi pribadi. Untuk mengatasi hal tersebut, dibangun sebuah sistem rekomendasi menggunakan pendekatan penyaringan berbasis konten dengan mengintegrasikan teknik pembobotan TF-IDF untuk menganalisis deskripsi produk serta perhitungan *cosine similarity* untuk mengukur tingkat kemiripan antar item. Hasil eksperimen dari penelitian pengguna menunjukkan bahwa sistem tersebut mampu mengurutkan rekomendasi produk secara relevan dari nilai kemiripan tertinggi hingga terendah dengan capaian nilai kemiripan tertinggi sebesar nol koma tujuh ratus dua puluh dua [7].

Penelitian sejalan lainnya yang mengkaji domain serupa dengan menerapkan metode *Content-Based Filtering* pada sistem rekomendasi pemilihan produk *skincare*. Penelitian ini dilatarbelakangi oleh tingginya variasi katalog produk di platform *e-commerce* kecantikan yang seringkali membuat konsumen kebingungan dalam menentukan pilihan yang sesuai dengan kondisi wajah. Berdasarkan pemodelan terhadap 30 sampel data uji, sistem berhasil memberikan rekomendasi yang presisi sesuai dengan studi kasus pencarian pengguna, dengan capaian nilai kemiripan tertinggi sebesar 0,61 pada produk rekomendasi teratas [8].

2.1. Sistem Rekomendasi

Sistem rekomendasi merupakan teknologi yang digunakan untuk memberikan saran atau rekomendasi kepada pengguna berdasarkan preferensi dan kebutuhan pengguna [9]. Sistem ini banyak digunakan dalam berbagai bidang seperti *e-commerce*, hiburan, dan layanan berbasis lokasi. Dengan memanfaatkan data pengguna dan karakteristik item, sistem rekomendasi mampu membantu pengguna dalam mengambil keputusan secara lebih efektif dan efisien.

2.2. Content-Based Filtering

Metode penyaringan berbasis konten atau *Content-Based Filtering* merupakan pendekatan sistem rekomendasi yang bekerja dengan cara

mencocokkan karakteristik item dengan kriteria preferensi pengguna. Dalam penerapannya mesin akan menganalisis profil fitur dari setiap item lalu mengukurnya secara matematis sehingga relevansi hasil pencarian sangat bergantung pada kelengkapan atribut data. Keunggulan utama dari metode ini adalah kemampuannya dalam menghasilkan rekomendasi yang sangat personal secara mandiri tanpa memerlukan riwayat interaksi dari pengguna lain [10].

2.3. Teks Preprocessing

Teks pra-pemrosesan merupakan tahapan awal yang sangat vital dalam pengolahan data tekstual guna memastikan kualitas masukan sebelum diproses oleh algoritma [11]. Dalam konteks sistem informasi, prosedur ini bertujuan untuk menghilangkan derau pada data mentah melalui serangkaian langkah sistematis seperti penyeragaman seluruh huruf menjadi huruf kecil, penghapusan karakter khusus yang tidak relevan, serta pemecahan kalimat menjadi unit kata mandiri atau tokenisasi.

2.4. Term Frequency–Inverse Document Frequency

TF-IDF merupakan metode pembobotan kata yang digunakan untuk mengukur tingkat kepentingan suatu kata dalam sebuah dokumen berdasarkan frekuensi kemunculannya serta jumlah dokumen yang memuat kata tersebut. Metode ini memberikan bobot lebih tinggi pada kata yang sering muncul dalam suatu dokumen tetapi jarang muncul pada dokumen lain [3].

Secara matematis, nilai TF-IDF diperoleh dari hasil perkalian antara Term Frequency (TF) dan Inverse Document Frequency (IDF). Nilai IDF dapat dirumuskan sebagai berikut:

$$IDF = \log \frac{N}{df} \quad (1)$$

di mana N adalah jumlah seluruh dokumen dan df adalah jumlah dokumen yang mengandung kata tersebut.

2.5. Cosine similarity

Cosine similarity merupakan metode yang digunakan untuk mengukur tingkat kemiripan antara dua objek yang direpresentasikan dalam bentuk vektor [12]. Nilai kemiripan dihitung berdasarkan cosinus sudut antara dua vektor, di mana semakin kecil sudut yang terbentuk, maka semakin tinggi tingkat kemiripannya. Secara matematis, *cosine similarity* dirumuskan sebagai berikut:

$$\text{cosine similarity} = \frac{A \cdot B}{|A||B|} \quad (2)$$

di mana A dan B merupakan dua vektor yang dibandingkan. Metode ini banyak digunakan dalam sistem rekomendasi untuk menghitung tingkat kemiripan antar item maupun antara profil pengguna dengan item, sehingga sistem dapat menghasilkan rekomendasi yang relevan sesuai dengan preferensi pengguna.

2.6. Precision

Metrik presisi merupakan salah satu parameter evaluasi utama dalam matriks kebingungan yang berfungsi untuk mengukur tingkat keakuratan sistem dalam memberikan rekomendasi kepada pengguna. Secara teknis nilai presisi merepresentasikan rasio antara jumlah item relevan yang berhasil ditemukan terhadap total keseluruhan item yang disarankan oleh mesin kepada pengguna pada saat proses pencarian berlangsung [13].

2.7. Recall

Recall merupakan parameter evaluasi yang sangat penting untuk mengukur kemampuan sistem dalam menemukan kembali seluruh informasi relevan dari dalam basis data [14]. Dalam sistem rekomendasi, nilai perolehan dihitung dengan membandingkan jumlah item relevan yang berhasil ditampilkan kepada pengguna terhadap total keseluruhan item relevan yang tersedia di dalam sistem tersebut.

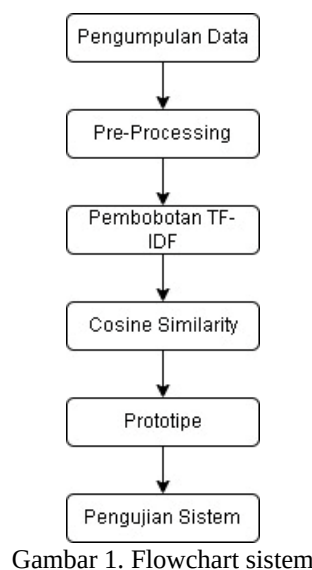
2.8. F1-Score

F1-Score merupakan metrik evaluasi yang menggabungkan nilai presisi dan perolehan ke dalam satu angka tunggal guna memberikan gambaran performa sistem secara keseluruhan. Penggunaan metrik ini sangat penting karena mampu memberikan nilai rata rata harmonik yang menyeimbangkan antara tingkat keakuratan dan kelengkapan hasil rekomendasi yang diberikan oleh algoritma [15].

3. METODE PENELITIAN

Metodologi penelitian yang digunakan mencakup rancangan alur penelitian, strategi pengumpulan data, tahapan perancangan sistem, penerapan algoritma penyaringan berbasis konten, hingga metode pengujian yang digunakan untuk mengukur tingkat keberhasilan aplikasi.

3.1. Alur Penelitian



Berdasarkan gambar 1, proses penelitian ini dijalankan melalui enam tahapan utama yang saling berkesinambungan. Langkah pertama diawali dengan tahap pengumpulan data untuk menghimpun seluruh informasi mentah mengenai atribut *coffee shop* di wilayah Surakarta. Setelah informasi terkumpul, proses dilanjutkan dengan tahap pra pemrosesan guna membersihkan dan menstandarkan format teks agar data siap untuk diolah ke dalam bahasa mesin.

Data yang telah seragam tersebut kemudian masuk ke tahap pembobotan menggunakan metode TF-IDF untuk menentukan nilai kepentingan dari setiap kata kunci fasilitas maupun menu. Nilai bobot yang dihasilkan selanjutnya diproses pada tahap kalkulasi *Cosine similarity* guna mengukur tingkat jarak kemiripan antara kriteria yang dicari oleh pengguna dengan data *coffee shop* yang tersedia di dalam basis data. Hasil perhitungan matematis tersebut kemudian diwujudkan pada tahap pembuatan prototipe berupa rancang bangun antarmuka aplikasi berbasis web. Sebagai langkah terakhir, penelitian ditutup dengan tahap pengujian sistem untuk mengevaluasi tingkat akurasi serta kelayakan perangkat lunak dalam menyajikan rekomendasi yang tepat sasaran.

3.2. Metode Pengumpulan Data

Metode pertama adalah studi literatur yang dilakukan dengan cara mengumpulkan serta mempelajari berbagai referensi tertulis dari jurnal ilmiah, buku teks, dan dokumen hasil penelitian sebelumnya yang berkaitan erat dengan tema sistem rekomendasi. Langkah ini sangat krusial untuk memperkuat landasan teori mengenai cara kerja algoritma serta memahami berbagai parameter yang umum digunakan dalam mengevaluasi performa sebuah sistem informasi.

Metode kedua adalah pengumpulan data melalui observasi digital untuk membangun basis data utama sistem. Data yang dihimpun dalam penelitian ini mencakup seratus *coffee shop* yang berlokasi di wilayah Surakarta. Informasi detail mengenai profil setiap *coffee shop* diperoleh melalui sumber daring terpercaya seperti layanan peta digital serta akun media sosial resmi dari masing masing tempat usaha tersebut. Atribut utama yang diekstraksi dari sumber data tersebut meliputi rentang harga produk, variasi kategori menu makanan atau minuman, serta ketersediaan fasilitas pendukung..

3.3. Perancangan Sistem

Perancangan sistem merupakan fase penerjemahan model matematis penyaringan berbasis konten ke dalam wujud perangkat lunak fungsional menggunakan kerangka kerja Laravel dan bahasa pemrograman PHP. Proses perancangan ini mencakup penyusunan arsitektur sistem yang mengintegrasikan basis data berisi seratus *coffee shop* di Surakarta dengan mesin pengolah rekomendasi di sisi. Secara teknis logika algoritma dibangun melalui skrip PHP untuk mengeksekusi rumus pembobotan dan kalkulasi

kemiripan secara otomatis setiap kali sistem menerima masukan kriteria.

3.4. Penerapan Content-Based Filtering

Bagian ini menguraikan tahapan inti dari algoritma penyaringan berbasis konten yang bertugas mencocokkan profil *coffee shop* dengan kriteria pencarian pengguna. Langkah pertama yang dilakukan mesin adalah membersihkan dan menyeragamkan format teks dari setiap atribut *coffee shop* seperti nama fasilitas dan kategori menu. Proses ini meliputi pengubahan seluruh huruf menjadi huruf kecil serta penghapusan karakter khusus maupun tanda baca agar bahasa mesin dapat membaca dan memproses data teks tersebut secara optimal sebelum masuk ke tahap perhitungan matematis.

Setelah seluruh teks seragam sistem akan menghitung bobot nilai dari setiap kata kunci menggunakan metode pembobotan frekuensi istilah dan *inversi frekuensi* dokumen. Dan tahap terakhir adalah mengukur tingkat kedekatan antara kriteria yang dimasukkan pengguna dengan data *coffee shop* menggunakan perhitungan geometri sudut ruang. Sistem akan menghitung nilai kemiripan kosinus dari dua vektor data untuk menghasilkan sebuah skor probabilitas.

3.5. Pengujian Sistem

Pengujian sistem berfungsi untuk mengukur tingkat keberhasilan dan keandalan prototipe aplikasi menggunakan metode matriks kebingungan. Evaluasi performa difokuskan pada tiga metrik utama yakni presisi untuk menilai persentase keakuratan antara saran yang diberikan dengan relevansi aslinya lalu perolehan untuk mengukur kemampuan algoritma dalam melacak seluruh *coffee shop* yang sesuai kriteria di dalam basis data serta skor f1 sebagai nilai rata rata harmonik penyeimbang dari kedua parameter tersebut. Melalui tahapan pengujian yang komprehensif dan

terukur ini diharapkan dapat ditarik sebuah kesimpulan objektif mengenai kelayakan perangkat lunak sebelum diimplementasikan secara nyata sebagai alat bantu pencarian *coffee shop* di Surakarta.

4. HASIL DAN PEMBAHASAN

Pada tahap ini dilakukan penerapan metode Content Based Filtering untuk merekomendasikan *coffee shop* di Surakarta. Sistem mengolah data berdasarkan tiga aspek penting, yaitu harga, pilihan menu, dan fasilitas.

4.1. Teks Preprocessing

Sebelum menghitung nilai kemiripan, seluruh data tekstual dari database *coffee shop* maupun preferensi pengguna wajib melalui tahap pra-pemrosesan teks. Langkah ini bertujuan untuk menyeragamkan format penulisan dan memecah string menjadi kata kunci mandiri agar proses pencocokan konten berjalan lebih akurat. Penelitian ini mengolah total 100 data *coffee shop* berdasarkan empat karakteristik utama, yaitu nama *coffee shop*, rata rata harga, menu, dan fasilitas. Tahapan pra-pemrosesan yang diterapkan diawali dengan *case folding* untuk mengubah seluruh huruf kapital di dalam aspek teks menjadi huruf kecil. Selanjutnya, dilakukan operasi eksklusi tanda baca guna menghapus simbol non-alfanumerik, termasuk menghilangkan tanda titik pada nominal harga agar bertransformasi menjadi bentuk numerik murni, serta tanda koma dan titik koma pemisah atribut. Proses tersebut diakhiri dengan *tokenizing*, yaitu memotong untaian string gabungan menjadi larik kata kunci tunggal yang siap digunakan pada tahap komputasi kemiripan. Sebagai contoh penerapan pada input pengguna, jika sistem menerima preferensi berupa: 20.000, Dessert, Makanan Berat AC, Wifi, Indoor. Berikut hasil transformasi data pada 10 sampel awal *coffee shop* disajikan secara terstruktur dalam Tabel 1.

Tabel 1. Hasil transformasi pra-pemrosesan

Kode	Nama	Data Mentah	Hasil
D1	Lucon Coffe	15.000 ; Non Coffee, Snack ; Ac, Wifi, Outdoor, Indoor	15000 non coffee snack ac wifi outdoor indoor
D2	Lana Coffe	25.000 ; Non Coffee, Snack, dessert, Makanan Berat ; Ac, Wifi, Indoor	25000 non coffee snack dessert makanan berat ac wifi indoor
D3	Satuperlima Coffe	20.000 ; Non Coffee, dessert, Makanan Berat ; Ac, Wifi, Indoor, Outdoor, Tempat Ibadah	20000 non coffee dessert makanan berat ac wifi indoor outdoor tempat ibadah
D4	ODDS	25.000 ; Non Coffee, Snack, dessert ; Ac, Wifi, Outdoor, Indoor	25000 non coffee snack dessert ac wifi outdoor indoor
D5	Getsee	25.000 ; Non Coffee, Snack, dessert ; Ac, Wifi, Outdoor, Indoor	25000 non coffee snack dessert ac wifi outdoor indoor
D6	Kaveloka	25.000 ; Non Coffee, Snack, dessert, Makanan Berat ; Ac, Wifi, Outdoor, Indoor	25000 non coffee snack dessert makanan berat ac wifi outdoor indoor
D7	Kopi Purwo	15.000 ; Non Coffe, Snack ; Wifi, Outdoor	15000 non coffe snack wifi outdoor
D8	The Gade Coffe	25.000 ; Non Coffee, Snack, dessert, Makanan Berat ; Ac, Wifi, Tempat Ibadah, Ruang Meeting	25000 non coffee snack dessert makanan berat ac wifi tempat ibadah ruang meeting
D9	Bercanda Coffee	20.000 ; Non Coffee ; Wifi, Outdoor	20000 non coffee wifi outdoor
D10	RND Coffee	25.000 ; Non Coffee, Snack, dessert ; Ac, Wifi, Outdoor, Indoor	25.000 non coffee snack dessert ac wifi outdoor indoor

Berdasarkan hasil yang ditunjukkan pada Tabel 1, proses prapemrosesan teks berhasil mengubah data mentah yang bervariasi menjadi kumpulan kata kunci yang seragam. Pada tahap penyeragaman huruf, seluruh karakter kapital diubah menjadi huruf kecil guna menghindari terjadinya duplikasi kata kunci yang sama akibat perbedaan penulisan. Selain itu, nilai numerik pada harga juga diselaraskan menjadi format angka nominal yang utuh agar memiliki struktur yang identik dengan preferensi yang dimasukkan oleh pengguna.

4.2. Pembobotan TF-IDF

Berdasarkan ekstraksi token terhadap 10 sampel *coffee shop* pada tahap sebelumnya, terdapat 13 atribut unik yang merepresentasikan karakteristik item. Nilai DF serta bobot IDF untuk masing-masing karakteristik tersebut tertera pada Tabel 2.

Tabel 2. Perhitungan parameter

Atribut	Nilai DF	Nilai IDF
15000	2	0.699
20000	2	0.699

Atribut	Nilai DF	Nilai IDF
25000	6	0.222
ac	8	0.097
dessert	7	0.155
indoor	7	0.155
makanan berat	4	0.398
non coffee	10	0.000
outdoor	8	0.097
Ruang meeting	1	1.000
snack	8	0.097
tempat ibadah	2	0.699
wifi	10	0.000

Data pada Tabel 2 menunjukkan bahwa atribut non coffee dan wifi menghasilkan nilai IDF sebesar 0,000 karena muncul pada seluruh sampel. Secara matematis, hal ini menandakan fitur tersebut bersifat umum sehingga tidak memiliki daya pembeda yang kuat untuk spesifikasi rekomendasi. Langkah berikutnya adalah mengalikan nilai keberadaan atribut pada setiap *coffee shop* dengan skor IDF untuk menghasilkan matriks pembobotan TF-IDF seperti yang disajikan dalam Tabel 3.

Tabel 3. Matriks pembobotan akhir

Kode	Coffee shop	Vektor Bobot
D1	Lucon coffe	15000: 0.699, ac: 0.097, indoor: 0.155, outdoor: 0.097, snack: 0.097
D2	Lana Coffe	25000: 0.22, ac: 0.097, dessert: 0.155, indoor: 0.155, makanan berat: 0.398, snack: 0.097
D3	Satuperlima Coffe	20000: 0.699, ac: 0.097 dessert: 0.155, indoor: 0.155, makanan berat: 0.398, outdoor: 0.097, tempat ibadah: 0.699
D4	ODDS	25000: 0.222; ac: 0.097; dessert: 0.155; indoor: 0.155; outdoor: 0.097; snack: 0.097
D5	Getsee	25000: 0.222, ac: 0.097, dessert: 0.155, indoor: 0.155, outdoor: 0.097, snack: 0.097
D6	Kaveloka	25000: 0.222, ac: 0.097, dessert: 0.155, indoor: 0.155 makanan berat: 0.398 outdoor: 0.097 snack: 0.097
D7	Kopi Purwo	15000: 0.699, outdoor: 0.097, snack: 0.097
D8	The Gade Coffe	25000: 0.222, ac: 0.097, dessert: 0.155, makanan berat: 0.398, ruang meeting: 1.000, snack: 0.097, tempat ibadah: 0.699
D9	Bercanda Coffee	20000: 0.699, outdoor: 0.097
D10	RND Coffee	25000: 0.222, ac: 0.097, dessert: 0.155, indoor: 0.155, outdoor: 0.097, snack: 0.097

Nilai bobot pada Tabel 3 menjadi vektor representasi tiap *coffee shop*. Vektor ini yang kemudian digunakan dalam perhitungan algoritma *Cosine Similarity* guna mengukur derajat kemiripan objek terhadap preferensi pencarian dari pengguna.

4.3. Komputasi Cosine similarity

Setelah nilai bobot TF-IDF diekstraksi ke dalam bentuk matriks, fase determinan dalam metode *Content-Based Filtering* adalah mengukur derajat kedekatan antara profil preferensi pengguna dengan entitas *coffee shop* pada *database*. Pengukuran ini diimplementasikan menggunakan metrik *cosine similarity*, yang mengkalkulasi nilai kosinus sudut di antara dua vektor multidimensi. Berdasarkan matriks pembobotan pada tabel 3 sebelumnya, berikut adalah proses kalkulasi skalar untuk panjang vektor profil pengguna dan salah satu sampel *coffee shop*, yaitu D3:

Panjang vektor pengguna =

$$user = \sqrt{0.699^2 + 0.155^2 + 0.398^2 + 0.097^2 + 0.155^2} = 0.839 \quad (3)$$

Panjang Vektor D3 =

$$D3 = \sqrt{0.699^2 + 0.155^2 + 0.398^2 + 0.097^2 + 0.155^2} = 1.097 \quad (4)$$

Perhitungan *Inversi frekuensi* =

$$(0.699 \times 0.699) + (0.097 \times 0.097) + (0.398 \times 0.398) + (0.155 \times 0.155) = 0.703 \quad (5)$$

Skor Akhir =

$$Similarity(user + D3) = \frac{0.703}{0.839 \times 1.097} = 0.764 \quad (6)$$

Tabel 4. Hasil rekomendasi

Peringkat	Coffee shop	Inversi frekuensi	Vektor	Skor
1	Satuperlima (D3)	0.703	1.097	0.764
2	Lana Coffe (D2)	0.216	0.524	0.491
3	ODDS (D4)	0.057	0.354	0.193
4	Getsee (D5)	0.057	0.354	0.193
5	Lucon (D1)	0.033	0.735	0.054

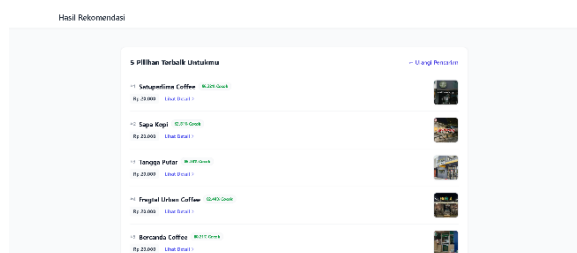
Hasil dari tabel 4 menunjukkan bahwa sistem berhasil mengidentifikasi dan memproyeksikan Satuperlima Coffe (D3) sebagai rekomendasi utama dengan skor kemiripan tertinggi 0.764. Tingginya skor ini berkorelasi langsung dengan tingkat presisi kecocokan pada atribut mayor berbobot tinggi, khususnya kriteria harga 20000 dan kategori menu makanan berat yang dicari oleh pengguna.

4.4. Prototipe

rancangan algoritma *content-based filtering* diimplementasikan ke dalam bentuk prototipe aplikasi berbasis web menggunakan *framework* laravel. Pembuatan prototipe ini bertujuan untuk memvisualisasikan bagaimana sistem menerima data masukan dari pengguna dan memprosesnya untuk menghasilkan rekomendasi yang sesuai.

Gambar 2. Halaman input preferensi

Gambar 2 menunjukkan tampilan halaman utama yang berfungsi sebagai media interaksi awal bagi pengguna untuk memasukkan kriteria pencarian *coffee shop*. Pada antarmuka ini, pengguna disediakan beberapa elemen formulir interaktif yang mencakup pilihan rentang harga, pilihan kategori menu, serta daftar fasilitas pendukung yang diinginkan. Pengguna dapat memilih opsi harga secara tunggal serta mencentang lebih dari satu pilihan untuk kategori menu dan fasilitas sesuai kebutuhan personal pengguna.



Gambar 3. Halaman Hasil Rekomendasi

Setelah pengguna menekan tombol proses pada halaman sebelumnya, sistem akan mengarahkan pengguna ke halaman hasil rekomendasi seperti yang diilustrasikan pada gambar 3. Halaman ini menyajikan daftar *coffee shop* di Surakarta yang telah diurutkan berdasarkan skor kemiripan tertinggi hasil kalkulasi algoritma *Cosine similarity*. Setiap *coffee shop* ditampilkan dalam bentuk kartu informasi yang memuat nama tempat, ringkasan fasilitas, serta nilai kecocokan dalam bentuk persentase atau angka desimal. Melalui visualisasi peringkat yang terstruktur ini, pengguna dapat dengan mudah mengidentifikasi *coffee shop* mana yang paling memenuhi kriteria pencarian pengguna untuk dikunjungi.

4.5. Pengujian Akurasi Sistem

Untuk membuktikan bahwa rekomendasi yang diberikan oleh sistem tepat sasaran, dilakukan evaluasi kinerja menggunakan metrik pengujian temu kembali informasi, yakni *Precision*, *Recall*, dan *F1-Score*. Guna memastikan konsistensi performa, pengujian dilakukan sebanyak 5 kali uji coba menggunakan variasi kata kunci preferensi yang berbeda. Pada setiap pengujian, sistem dikonfigurasi untuk menampilkan 5 daftar rekomendasi teratas.

Sebagai contoh pemodelan matematis pada skenario pengujian pertama, pengguna memasukkan preferensi pencarian berupa fasilitas wifi, outdoor, dan harga 15.000. Setelah divalidasi, dari 5 rekomendasi yang dikeluarkan sistem, terdapat 4 *coffee shop* yang benar-benar relevan dengan kriteria pengguna, sementara total keseluruhan *coffee shop* yang memiliki kriteria tersebut di dalam basis data berjumlah 5. Perhitungan matematis akurasinya adalah sebagai berikut:

$$Precision = \frac{4}{5} \times 100\% = 80\% \quad (7)$$

Melalui perhitungan ini, diperoleh nilai precision sebesar 80% yang berarti mayoritas rekomendasi yang disajikan sistem sudah sesuai dengan kriteria asli yang dicari oleh pengguna.

$$Recall = \frac{4}{5} \times 100\% = 80\% \quad (8)$$

Nilai perolehan sebesar 80% ini membuktikan kemampuan sistem yang mumpuni dalam melacak dan menampilkan alternatif *coffee shop* yang cocok dari keseluruhan isi basis data.

F1-Score =

$$F1 - Score = 2 \times \frac{0.80 \times 0.80}{0.80 + 0.80} \times 100\% = 80\% \quad (9)$$

Skor *F1-Score* sebesar 80% menunjukkan keseimbangan performa yang sangat baik antara ketepatan rekomendasi dan kemampuan pelacakan sistem.

Proses evaluasi yang sama diterapkan pada keempat skenario pengujian lainnya secara mandiri. Rekapitulasi hasil komputasi dari 5 kali simulasi pencarian dijabarkan pada Tabel 5.

Tabel 5. Rangkuman hasil evaluasi sistem

No	Skenario Uji	Metrik	Hasil
Uji 1	Wifi,outdoor,15000	Precision	80.00
		Recall	80.00
		F1-Score	80.00
Uji 2	Ac,wifi,indoor,makanan berat,25000	Precision	100
		Recall	80.00
		F1-Score	80.00
Uji 3	dessert,ac,wifi, indoor,20000	Precision	80.00
		Recall	100
		F1-Score	88.89
Uji 4	snack,dessert,ac, wifi, outdoor,25000	Precision	100
		Recall	100
		F1-Score	100
Uji 5	non coffee,snack,15000	Precision	80.00
		Recall	80.00
		F1-Score	80.00

Berdasarkan hasil evaluasi di tabel 5, kalkulasi menunjukkan bahwa nilai performa sistem telah memenuhi dan melampaui target standar evaluasi. Perolehan nilai rata-rata *precision* mencapai 88,00%, rata-rata *recall* berada di angka 88,67%, dan rata-rata *f1-score* menyentuh angka 87,96%. Pencapaian skor tinggi yang merata di atas 80% pada seluruh metrik ini memperkuat kesimpulan bahwa algoritma *content-based filtering* yang diterapkan memiliki keandalan yang tinggi, sehingga sangat layak untuk digunakan sebagai alat bantu cerdas dalam pemilihan *coffee shop* di Surakarta.

5. KESIMPULAN DAN SARAN

Penelitian ini secara keseluruhan menyimpulkan bahwa sistem rekomendasi *coffee shop* di Surakarta berhasil dirancang dan diimplementasikan dengan baik melalui perpaduan kerangka kerja Laravel dan algoritma penyaringan berbasis konten *Content-Based Filtering*. Penerapan algoritma ini terbukti efektif, yang ditunjukkan melalui perolehan nilai kemiripan tertinggi sebesar 0,764 pada objek rekomendasi Satuperlima Coffe. Selain itu, hasil evaluasi performa membuktikan bahwa prototipe aplikasi ini memiliki tingkat keandalan yang sangat memuaskan dengan perolehan nilai rata-rata *Precision* sebesar 88,00%, *Recall* sebesar 88,67%, dan rata-rata *F1-Score* mencapai 87,96%. Sebagai langkah penyempurnaan dan saran untuk pengembangan sistem di masa mendatang, peneliti selanjutnya diharapkan dapat memperluas skala basis data *coffee shop* serta mencoba mengintegrasikan algoritma saat ini dengan pendekatan penyaringan *Collaborative Filtering*, agar hasil rekomendasi yang disajikan menjadi jauh lebih presisi, dinamis, dan bervariasi.

DAFTAR PUSTAKA

- [1] D. Warta, P. Pramono, and J. Maulindar, "Sistem Rekomendasi Kuliner Ikonik Kota Solo Menggunakan Metode Content Based Filtering," *Jurnal Teknologi Terpadu*, vol. 11, no. 1, pp. 78–86, 2025.
- [2] L. H. Aljihadu, "SISTEM REKOMENDASI WISATA KULINER DI GUNUNGKIDUL MENGGUNAKAN METODE CONTENT BASED FILTERING," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5955.
- [3] V. Meida Hersianty, E. Larasati Amalia, D. Puspitasari, and D. Wahyu Wibowo, "PENERAPAN ALGORITMA TF-IDF DAN COSINE SIMILARITY DALAM SISTEM REKOMENDASI LOWONGAN PEKERJAAN," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 1619–1625, Jan. 2025, doi: 10.36040/jati.v9i1.12406.
- [4] R. Oktavian and F. Amin, "Perancangan Sistem Rekomendasi Laptop dengan Model Prototyping dan Penerapan Content-Based Filtering Approach pada ELS Computer Shop Semarang," *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 1, pp. 66–80, 2024.
- [5] R. Ardiansyah, M. Ari Bianto, and B. D. Saputra, "Sistem Rekomendasi Buku Perpustakaan Sekolah menggunakan Metode Content-Based Filtering," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 2, pp. 510–518, Oct. 2023, doi: 10.37859/coscitech.v4i2.5131.
- [6] B. C. Putro, D. Remawati, S. H. Fitriasih, and K. Sandradewi, "SISTEM REKOMENDASI COFFEE SHOP DI KOTA SURAKARTA DENGAN KATEGORI FASILITAS MENGGUNAKAN ITEM-BASED COLLABORATIVE FILTERING," *JRIS: Jurnal Rekayasa Informasi Swadharma*, vol. 6, no. 1, pp. 93–103, 2026.
- [7] N. Azizah and A. F. Rozi, "Sistem Rekomendasi Produk Somethinc Menggunakan Metode Content-based Filtering," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 3, pp. 461–468, 2024.
- [8] A. Sulami, V. Atina, and N. Nurmalitasari, "Penerapan Metode Content Based Filtering dalam Sistem Rekomendasi Pemilihan Produk Skincare," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 9, no. 2, p. 172, Dec. 2024, doi: 10.30998/string.v9i2.24066.
- [9] Crismastiana Koloman, Raihan Maulana, Raisya Dwi Zahra Putri, and Wahyu Abadi Harahap, "Sistem Rekomendasi Pekerjaan di bidang IT Menggunakan Algoritma Content-Based Filtering," *Journal of Creative Student Research*, vol. 1, no. 6, pp. 78–88, Dec. 2023, doi: 10.55606/jcsrpolitama.v1i6.2992.
- [10] B. M. Kumarahadi, Y. K. Kumarahadi, and D. Ridhwanullah, "Sistem Rekomendasi Makanan Kucing Menggunakan Metode Content-Based Filtering," *JURNAL KRIDATAMA SAINS DAN TEKNOLOGI*, vol. 7, no. 01, pp. 46–61, Feb. 2025, doi: 10.53863/kst.v7i01.1471.

- [11] S. Rahmadhani, L. Hakim, G. H. Wibowo, S. P. Kristanto, and E. M. Rini, "Sistem Rekomendasi Penelusuran Buku Berbasis Content-Based Filtering dengan Pembobotan TF-RF," *Jurnal Informatika Polinema*, vol. 10, no. 4, pp. 491–500, 2024.
- [12] T. Ridwansyah, B. Subartini, and S. Sylviani, "Penerapan Metode Content-Based Filtering pada Sistem Rekomendasi," *Math. Sci. Appl. J.*, vol. 4, no. 2, pp. 70–77, 2024.
- [13] M. Ridho Robbi and A. Voutama, "EVALUASI PERFORMA ALGORITMA SIMILARITY PADA DATASET FILM INDONESIA MENGGUNAKAN METRIK PRECISION DAN RECALL," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 10, no. 2, pp. 3623–3627, Mar. 2026, doi: 10.36040/jati.v10i2.18250.
- [14] S. A. Siti Aminah, "Perbandingan Keakuratan Sistem Rekomendasi Produk Berbasis Content-Based Filtering Dan Collaborative Filtering Pada E-Commerce Shopee Menggunakan Matrik Precision, Recall Dan F1-Score," *Online Repository of Universitas NU Kalimantan Selatan*, pp. 1–18, 2024.
- [15] M. Yazid, D. A. Sani, and N. M. Anggadimas, "Sistem Rekomendasi Penerima Bantuan Sosial APBD Menggunakan Metode Support Vector Machine," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 6, no. 1, pp. 54–61, Apr. 2024, doi: 10.28926/ilkomnika.v6i1.619