

Penerapan Content-Based Filtering untuk Rekomendasi UMKM Kuliner di Kota Surakarta

Elisabeth Puteri Sekar Arum^{1*}, Sofiana Dewi Larasati², Alyesa Putri Aprilia³, Abdillah Ichsan⁴, Aditya Putra Mahendra⁵

¹ Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
^{1*} 230103019@mhs.udb.ac.id

² Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
² 230103019@mhs.udb.ac.id

³ Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
³ 230103007@mhs.udb.ac.id

⁴ Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
⁴ 230103001@mhs.udb.ac.id

⁵ Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta
⁵ 230103002@mhs.udb.ac.id

Abstrak— Kota Surakarta terkenal sebagai salah satu destinasi wisata kuliner legendaris yang memiliki keanekaragaman makanan tradisional yang kaya. Pertumbuhan Usaha Mikro, Kecil, dan Menengah (UMKM) kuliner yang sangat pesat sering kali menimbulkan tantangan berupa limpahan informasi bagi wisatawan maupun warga lokal dalam memilih tempat makan yang sesuai dengan preferensi selera dan kenyamanan mereka. Penelitian ini bertujuan mengembangkan sistem rekomendasi kuliner berbasis Content-Based Filtering (CBF) yang memanfaatkan pencocokan multi-atribut konten yang mencakup kategori utama, menu andalan, kata kunci ulasan, dan fasilitas fisik tempat makan. Data primer dikumpulkan langsung dari Google Maps tanpa melibatkan rating subjektif untuk menghindari bias penilaian. Sistem ini menggunakan algoritma Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengekstrak dan memberikan bobot pada fitur konten serta algoritma Cosine Similarity untuk menghitung indeks kemiripan antar-UMKM kuliner berdasarkan preferensi fasilitas dan menu pengguna. Hasil eksperimen komputasi menunjukkan bahwa pemodelan CBF berbasis atribut fasilitas mampu menghasilkan rekomendasi kuliner yang sangat presisi, relevan, dan andal. Berdasarkan simulasi pengujian query menu soto sapi dengan fasilitas parkir, sistem merekomendasikan Soto Arinda (Kerten) sebagai rekomendasi utama dengan nilai similarity sebesar 0,949, sekaligus menjadi instrumen strategis dalam memperluas jangkauan promosi digital bagi pelaku UMKM kuliner di Kota Surakarta.

Kata kunci— Content-Based Filtering, Google Maps, Kuliner Surakarta, Fasilitas, TF-IDF.

Abstract— Surakarta City is widely recognized as a legendary culinary tourism destination with a rich diversity of traditional foods. The rapid growth of culinary Micro, Small, and Medium Enterprises (MSMEs) often presents an information overload challenge for both tourists and local residents in selecting dining venues that match their taste and convenience preferences. This study aims to develop a culinary recommendation system based on Content-Based Filtering (CBF) by leveraging multi-attribute content matching, including main categories, signature menus, review keywords, and physical dining facilities. Primary data was crawled directly from Google Maps without involving subjective ratings to avoid evaluation bias. The system utilizes the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm to extract and weight content features and the Cosine Similarity algorithm to compute the similarity index between culinary MSMEs based on users' facility and menu preferences. The computational experimental results demonstrate that CBF modeling based on facility attributes delivers highly precise, relevant, and reliable culinary recommendations. Based on the query testing simulation for a beef soto menu with parking facilities, the system recommends Soto Arinda (Kerten) as the primary suggestion with a similarity value of 0.949, serving as a strategic instrument to expand digital promotion for culinary MSMEs in Surakarta.

Keywords— Content-Based Filtering, Google Maps, Surakarta Culinary, Facilities, TF-IDF

I. PENDAHULUAN

Sektor kuliner merupakan salah satu pilar utama penggerak perekonomian kreatif dan daya tarik pariwisata di Kota Surakarta. [1]

Sebagai kota yang sarat akan nilai budaya, Surakarta menawarkan berbagai macam kuliner

ikonik seperti Selat Solo, Timlo, Soto Sapi, Tengkleng, hingga jajanan tradisional khas seperti Tahok dan Dawet Telasih. Meningkatnya aktivitas digital mendorong konsumen untuk mencari referensi kuliner melalui internet.[2] Namun, keberadaan ratusan

UMKM kuliner di Surakarta justru memicu terjadinya fenomena *information overload*, di mana konsumen merasa kesulitan dan membutuhkan waktu lama untuk menyaring informasi guna menemukan tempat makan yang sesuai dengan preferensi pribadi mereka. [2]

Secara konvensional, pencarian kuliner masih mengandalkan mesin pencari umum atau rekomendasi mulut ke mulut yang tidak personal.[2] Guna mengatasi keterbatasan ini, teknologi sistem rekomendasi diterapkan untuk menyaring data secara otomatis berdasarkan ketertarikan spesifik pengguna.[3] Pendekatan utama dalam sistem rekomendasi meliputi Collaborative Filtering (CF) dan Content-Based Filtering (CBF).[4] Pendekatan CF merekomendasikan item berdasarkan rekam jejak penilaian dari pengguna lain yang memiliki kesamaan minat. Meskipun efektif, CF memiliki kelemahan struktural berupa masalah cold-start dan keterjarangan data (data sparsity), di mana sistem gagal memberikan rekomendasi yang akurat untuk kuliner baru yang belum memiliki rating dari pengguna, atau ketika basis data pengguna masih sangat sedikit.[5]

Metode Content-Based Filtering (CBF) menawarkan keunggulan berupa kemandirian terhadap pengguna lain (user independence). CBF bekerja dengan menganalisis deskripsi fitur dari objek yang diminati oleh pengguna, lalu mencari objek lain yang memiliki kemiripan karakteristik tertinggi.[6] Karakteristik kuliner tradisional di Surakarta sangat kaya akan atribut tekstual seperti nama menu, cita rasa dominan, dan fasilitas penunjang tempat makan (seperti area parkir, akses WiFi, toilet, dan mushola), sehingga pemodelan berbasis konten menjadi pilihan yang sangat relevan dan logis. [7]

Dalam pengembangan sistem rekomendasi ini, performa algoritma CBF sangat dipengaruhi oleh representasi teks yang digunakan.[8] Oleh karena itu, penelitian ini mengintegrasikan metode pembobotan kata Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengekstrak bobot penting dari setiap kata kunci kuliner dan fasilitas, yang kemudian dipadukan dengan algoritma Cosine Similarity

untuk menghitung kedekatan sudut antar-vektor profil kuliner.[9]

Kebaruan utama dalam penelitian ini terletak pada proses akuisisi dan kurasi data yang dilakukan.[10] Data dikumpulkan secara langsung dari platform Google Maps, yang merupakan direktori spasial dan bisnis paling aktif digunakan oleh masyarakat.[11] Guna memberikan rekomendasi yang objektif dan bebas dari bias penilaian, penelitian ini memfokuskan analisis pembahasan pada pencocokan preferensi fasilitas fisik tempat makan dan menu andalan.[11] Fasilitas seperti area parkir, toilet, mushola, AC, dan layanan bawa pulang (take away) merepresentasi kebutuhan kenyamanan riil konsumen saat berkunjung. Penerapan fitur fasilitas ini diharapkan mampu menghadirkan luaran rekomendasi yang sangat sesuai dengan profil kebutuhan fisik pengguna sekaligus membantu digitalisasi profil pelaku UMKM kuliner di Kota Surakarta.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental yang diintegrasikan dengan pemrosesan bahasa alami (Natural Language Processing) untuk membangun sistem rekomendasi. Tahapan penelitian dilakukan secara terstruktur yang meliputi akuisisi data, pra-pemrosesan teks, representasi fitur kata dengan TF-IDF, pengukuran kemiripan dengan Cosine Similarity.

A. Akuisisi Data Google Maps

Tahap awal penelitian adalah proses pengumpulan data spesifik direktori bisnis UMKM kuliner dari platform Google Maps di wilayah Surakarta. Akuisisi data dilakukan menggunakan metode ekstraksi informasi (web scraping) secara otomatis. Struktur atribut data yang ditarik untuk membangun basis data profil meliputi nama UMKM kuliner, kategori bisnis, menu makanan andalan, kata kunci ulasan (review tags), dan kelengkapan fasilitas fisik tempat makan.

B. Pra-pemrosesan Teks

Sebelum data teks diproses oleh algoritma pembobotan, seluruh string teks pada atribut konten digabungkan dan dilewatkan pada tahap pra-pemrosesan bahasa alami (*natural language preprocessing*). Tahapan ini bertujuan mereduksi derau teks (*textual noise*) melalui langkah-langkah berikut :

1. *Case Folding*: Mengonversi seluruh huruf kapital dalam dokumen teks menjadi huruf kecil (*lowercase*) agar pencocokan kata bersifat konsisten.
2. *Data Cleaning*: Menghapus elemen non-alfabet seperti angka, tanda baca, emoji, dan karakter khusus lainnya.
3. *Tokenization*: Memecah kalimat atau frase deskripsi kuliner menjadi urutan kata tunggal (*token*).
4. *Stopword Removal*: Mengeliminasi kata-kata fungsional yang tidak memiliki bobot semantik penting (seperti "yang", "di", "dan", "untuk", "dengan") menggunakan daftar stopwords bahasa Indonesia standar.
5. *Stemming*: Mengubah kata berimbuhan menjadi kata dasarnya (misalnya "menyajikan" menjadi "saji") guna menyeragamkan variasi morfologi kata.

C. Pembobotan Nilai TF-IDF

Kata-kata unik hasil pemrosesan teks ditransformasikan menjadi representasi vektor numerik menggunakan algoritma pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Algoritma ini berjalan melalui empat langkah perhitungan matematis berikut :

1. *Term Frequency (TF)*:
Mengukur frekuensi kemunculan term t dalam suatu dokumen d .
2. *Document Frequency (DF)*:
Menghitung jumlah total dokumen dalam korpus yang mengandung term t .
3. *Inverse Document Frequency (IDF)*:
Mengukur tingkat keunikan term berdasarkan sebarannya di seluruh korpus N . Formulasi perhitungan IDF dinyatakan pada Persamaan (1):

$$IDF_t = \log_{10} \left(\frac{N}{df_t} \right) + 1$$

4. *Kalkulasi Bobot TF-IDF (W)*:

Mengalikan nilai frekuensi lokal dengan bobot global kata. Formulasi akhir pembobotan dinyatakan pada Persamaan (2):

$$w_{t,d} = tf_{t,d} \times IDF_t$$

D. Pengukuran Kemiripan dengan Cosine Similarity

Tahap akhir adalah mengukur tingkat kemiripan arah vektor antara input preferensi pengguna (vektor query A) dengan representasi profil UMKM kuliner dalam basis data (vektor dokumen B). Algoritma *Cosine Similarity* menghitung nilai kosinus dari sudut yang terbentuk di antara kedua vektor tersebut. Formulasi matematis dari fungsi kesamaan kosinus dinyatakan pada Persamaan (3):

$$\text{CosineSimilarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Keterangan:

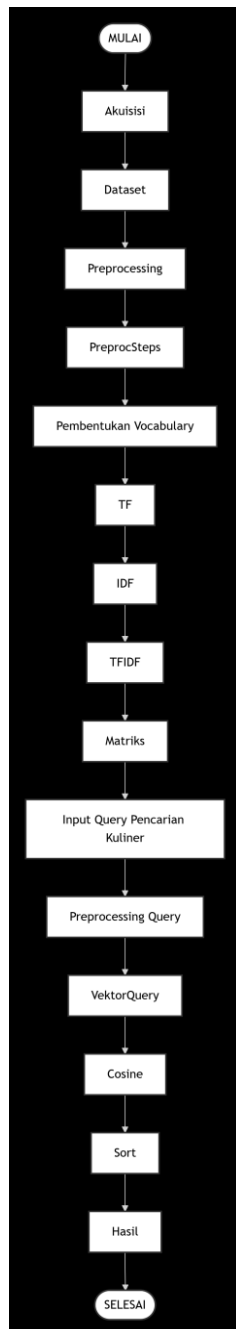
- $A \cdot B$: Hasil perkalian titik (*dot product*) antara vektor A dan vektor B .
- $\|A\|$: Magnitudo atau panjang dari vektor A .
- $\|B\|$: Magnitudo atau panjang dari vektor B .
- n : Jumlah total dimensi fitur unik dalam ruang vektor.

Nilai kemiripan berkisar antara 0 (tidak ada kesamaan sama sekali) hingga 1 (kesamaan mutlak atau identik).¹ Nilai kemiripan yang diperoleh kemudian digunakan untuk melakukan pengurutan (*sorting*) guna menampilkan daftar UMKM kuliner terbaik kepada pengguna.²

III. HASIL DAN PEMBAHASAN

A. Diagram Alir Pemodelan Sistem

Untuk memudahkan visualisasi tahapan logis pemrosesan informasi dari input preferensi pengguna hingga keluaran rekomendasi, dirancang suatu diagram alir pemodelan sistem rekomendasi Content-Based Filtering (CBF) seperti yang diilustrasikan pada Gambar 1.



Gambar 1 Diagram Alir Pemodelan Sistem

B. Akuisisi Data Google Maps

Proses ekstraksi data dari platform Google Maps berhasil menghimpun profile lengkap dari 50 entitas UMKM kuliner di Kota Surakarta. Atribut fasilitas fisik tempat makan (seperti parkir, toilet, mushola, WiFi, AC) diikutsertakan sebagai fitur konten utama bersanding dengan kategori dan menu andalan. Sampel data profil fisik UMKM kuliner Surakarta hasil akuisisi disajikan pada Tabel I.

Tabel 1. Sampel Profil Data Kuliner Surakarta dan Fasilitas Fisik dari Google Maps

Nama Kuliner	Kategori Utama Maps	Menu Andalan	Keyword Ulasan	Fasilitas
Soto Arinda	Warung Soto	Soto Sapi, Paru Goreng, Gorengan Garing	Langganan Warga, Kuah Bening Segar, Harga Bersahabat	Parkir, Makan di Tempat, Take Away
Timlo Sastro	Restoran Sup Jawa	Timlo Komplit, Sosis Solo Goreng	Ramai, Sarapan Legendaris, Kuah Bening Otentik	Parkir, Toilet, Makan di Tempat
Selat Solo Mbak Lies	Restoran Masakan Jawa	Selat Galantin, Sup Matahari	Unik, Interior Keramik Antik, Manis Gurih	WiFi, Parkir, Toilet
Tengkleng Rica Pak Manto	Restoran Olahan Kambing	Tengkleng Rica, Sate Buntel, Sumsum	Pedas Mantap, Kaya Rempah, Daging Empuk, Antre	Parkir, Mushola, Makan di Tempat
Tahok Pak Citro	Jajanan Tradisional	Tahok Kuah Jahe, Kembang Tahu	Lembut Sutra, Hangat Jahe, Legendaris Sejak 1958	Parkir, Mushola, Makan di Tempat
Sate Buntel Badran	Restoran Sate Modern	Sate Buntel Lembut, Nasi Goreng	Rating Tinggi, Inovasi Kuliner, Juicy	Parkir, Ruang Keluarga, Toilet

Sebagai basis pemodelan komputasi simulasi, dibentuk korpus dokumen utama ($N = 4$) yang merepresentasikan variasi karakteristik menu dan fasilitas yang berbeda seperti yang disajikan pada Tabel II.

Tabel 2 Korpus Dokumen Komputasi Berdasarkan Fitur Menu dan Fasilitas

ID	Nama Kulin er	Kategori Utama Maps	Menu Andalan	Keyword Ulasan	Fasilitas	Nama Kulin er
D 1	Soto Arinda	Warung Soto	Soto Sapi, Paru Goreng, Gorengan Garing	Langganan, Warga, Kuah Bening Segar, Harga Bersahabat	Parkir, Makan di Tempat, Take Away	Soto Arinda
D 2	Selat Solo Mbak Lies	Restoran Masakan Jawa	Selat Galantin, Sup Matahari	Unik, Interior Keramik Antik, Manis Guruh	WiFi, Parkir, Toilet	Selat Solo Mbak Lies
D 3	Tahok Pak Citro	Jajanan Tradisional	Tahok Kuah Jahe, Kembang Tahu	Lembut Sutra, Hangat Jahe, Legendaris Sejak 1958	Parkir, Mushola, Makan di Tempat	Tahok Pak Citro
D 4	Sate Buntel Badran	Restoran Sate Modern	Sate Buntel, Lembut, Nasi Goreng	Rating Tinggi, Inovasi Kuliner, Juicy	Parkir, Ruang Keluar, Toilet	Sate Buntel Badran

C. Proses Pra-pemrosesan Teks

Gabungan teks deskripsi fitur dari korpus Tabel II diproses melalui 5 tahap NLP. Guna menunjukkan mekanisme pemrosesan secara transparan, disajikan simulasi transformasi teks riil pada dokumen **Soto Arinda (Kerten) (D1)** :

1. Teks Mentah Gabungan: *"Warung Soto, Soto Sapi, Paru Goreng, Gorengan Garing, Langganan Warga, Kuah Bening Segar, Harga Bersahabat, Parkir, Makan di Tempat, Take Away"*
2. Hasil Case Folding: *"warung soto, soto sapi, paru goreng, gorengan garing, langganan warga, kuah bening segar, harga bersahabat, parkir, makan di tempat, take away"*
3. Hasil Data Cleaning: *"warung soto soto sapi paru goreng gorengan garing langganan warga kuah bening segar harga bersahabat parkir makan di tempat take away"*
4. Hasil Tokenization: ['warung', 'soto', 'soto', 'sapi', 'paru', 'goreng', 'gorengan', 'garing', 'langganan', 'warga', 'kuah', 'bening', 'segar', 'harga', 'bersahabat',

'parkir', 'makan', 'di', 'tempat', 'take', 'away']

5. Hasil Stopword Removal (menghilangkan kata non-informasi "di", "take", "away"): ['warung', 'soto', 'soto', 'sapi', 'paru', 'goreng', 'gorengan', 'garing', 'langganan', 'warga', 'kuah', 'bening', 'segar', 'harga', 'bersahabat', 'parkir', 'makan', 'tempat']
6. Hasil Stemming (mengubah kata berimbuhan "gorengan" -> "goreng", "langganan" -> "langgan", "bersahabat" -> "sahabat"): ['warung', 'soto', 'soto', 'sapi', 'paru', 'goreng', 'goreng', 'garing', 'langgan', 'warga', 'kuah', 'bening', 'segar', 'harga', 'sahabat', 'parkir', 'makan', 'tempat']

D. Perhitungan Pembobotan Nilai TF-IDF

Pada langkah ini, seluruh term hasil pemrosesan teks ditransformasikan secara matematis menggunakan algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Pengujian dilakukan dengan memodelkan preferensi pencarian pengguna melalui kalimat query berikut :

Query (Q): *"Ingin makan soto sapi yang kuah bening segar dan ada fasilitas parkir"*

Proses kalkulasi nilai bobot dibagi menjadi empat langkah operasional yang sistematis sebagai berikut:

1) Perhitungan Term Frequency (TF)

Term Frequency (TF) dihitung dengan mengukur frekuensi kemunculan dari setiap term yang terdapat pada vocabulary utama terhadap masing-masing dokumen dan query Q. Vocabulary kata kunci yang dievaluasi difokuskan pada: {soto, sapi, kuah, bening, segar, parkir, wifi}. Hasil ekstraksi nilai TF disajikan pada Tabel III.

Tabel 3 Matriks Frekuensi Kemunculan Term (Term Frequency - TF)

term	Tf di q	Tf di D1	Tf di D2	Tf di D3	Tf di D4
Soto	1	2	0	0	0
Sapi	1	1	0	0	0

Kuah	1	1	0	1	0
Bening	1	1	0	0	0
Segar	1	1	0	0	0
Parkir	1	1	1	1	1
Wifi	0	0	1	0	0

2) Perhitungan Document Frequency (DF)

Document Frequency (DF) dihitung dengan mencari jumlah dokumen dalam korpus yang mengandung term t minimal satu kali. Berdasarkan korpus $N = 4$ pada Tabel II, nilai DF untuk masing-masing term adalah sebagai berikut :

- Term "soto", "sapi", "bening", "segar", dan "wifi" masing-masing hanya muncul di 1 dokumen, sehingga nilai $df = 1$.
- Term "kuah" muncul di dokumen D1 dan D3, sehingga nilai $df = 2$.
- Term "parkir" muncul di seluruh dokumen (D1, D2, D3, D4), sehingga nilai $df = 4$.

3) Perhitungan Inverse Document Frequency (IDF)

Inverse Document Frequency (IDF) mengukur tingkat keunikan atau bobot penting dari suatu term di dalam seluruh korpus. Nilai IDF dihitung dengan mengaplikasikan fungsi logaritma pada pembagian total dokumen N dengan nilai DF masing-masing term. Formulasi perhitungan IDF menggunakan Persamaan (1) dan disajikan secara lengkap pada Tabel IV.

Tabel 4 Hasil Perhitungan Nilai DF dan IDF untuk Term Utama

Term (t)	Kemunculan Term pada Dokumen	Document Frequency (df _t)	Formula IDF (IDF _t = $\log_{10}(4 / df_t + 1)$)	Nilai IDF
soto	D1	1	$\log_{10}(4/1) + 1 = 0.602 + 1$	1.602
sapi	D1	1	$\log_{10}(4/1) + 1 = 0.602 + 1$	1.602
kuah	D1, D3	2	$\log_{10}(4/2) + 1 = 0.301 + 1$	1.301
bening	D1	1	$\log_{10}(4/1) + 1 = 0.602 + 1$	1.602
segar	D1	1	$\log_{10}(4/1) + 1 = 0.602 + 1$	1.602
parkir	D1, D2, D3, D4	4	$\log_{10}(4/4) + 1 = 0.000 + 1$	1.000
wifi	D2	1	$\log_{10}(4/1) + 1 = 0.602 + 1$	1.602

4) Perhitungan Bobot Akhir TF-IDF (W)

Langkah terakhir dalam pembobotan adalah mengalikan nilai frekuensi lokal (TF) pada Tabel III dengan bobot kepentingan global (IDF) pada Tabel IV menggunakan Persamaan (2). Hasil perkalian ini membentuk bobot akhir TF-IDF (W) yang disajikan pada Tabel V.

Tabel 5 Matriks Pembobotan TF-IDF untuk Dokumen dan Query

Term	IDF	Q	D1	D2	D3	D4
Soto	1.602	1.602	3.204	0.000	0.000	0.000
Sapi	1.602	1.602	1.602	0.000	0.000	0.000
Kuah	1.301	1.301	1.301	0.000	1.301	0.000
Bening	1.602	1.602	1.602	0.000	0.000	0.000
Segar	1.602	1.602	1.602	0.000	0.000	0.000
Parkir	1.000	1.000	1.000	1.000	1.000	1.000
Wifi	1.602	0.000	0.000	1.602	0.000	0.000

Berdasarkan Tabel V, bentuk representasi arah vektor numerik didefinisikan sebagai berikut :

$$\vec{V}_Q =$$

- [1.602, 1.602, 1.301, 1.602, 1.602, 1.000, 0.000]

$$\vec{V}_{D1} =$$

- [3.204, 1.602, 1.301, 1.602, 1.602, 1.000, 0.000]

$$\vec{V}_{D2} =$$

- [0.000, 0.000, 0.000, 0.000, 0.000, 1.000, 1.602]

$$\vec{V}_{D3} =$$

- [0.000, 0.000, 1.301, 0.000, 0.000, 1.000, 0.000]

$$\vec{V}_{D4} =$$

- [0.000, 0.000, 0.000, 0.000, 0.000, 1.000, 0.000]

E. Perhitungan Cosine Similarity dan Hasil Rekomendasi

Arah kesamaan vektor dihitung secara sistematis menggunakan formula kesamaan kosinus sesuai Persamaan (3).

1) Perhitungan Euclidean Norm (Magnitudo Vektor)

$$\begin{aligned}\|\vec{V}_Q\| &= \sqrt{1.602^2 + 1.602^2 + 1.301^2 + 1.602^2 + 1.602^2 + 1.000^2} = \sqrt{12.957} \approx 3.600 \\ \|\vec{V}_{D1}\| &= \sqrt{3.204^2 + 1.602^2 + 1.301^2 + 1.602^2 + 1.602^2 + 1.000^2} = \sqrt{20.657} \approx 4.545 \\ \|\vec{V}_{D2}\| &= \sqrt{1.000^2 + 1.602^2} = \sqrt{3.566} \approx 1.888 \\ \|\vec{V}_{D3}\| &= \sqrt{1.301^2 + 1.000^2} = \sqrt{2.693} \approx 1.641 \\ \|\vec{V}_{D4}\| &= \sqrt{1.000^2} = 1.000\end{aligned}$$

2) Perhitungan Dot Product dan Cosine Similarity

Kesesuaian dengan D_1 (Soto Arinda (Kerten)):

$$\begin{aligned}\vec{V}_Q \cdot \vec{V}_{D1} &= (1.602 \times 3.204) + (1.602 \times 1.602) + (1.301 \times 1.301) + \\ &+ (1.602 \times 1.602) + (1.602 \times 1.602) + (1.000 \times 1.000) \\ &= 15.524\end{aligned}$$

$$\text{Sim}(Q, D_1) = \frac{15.524}{3.600 \times 4.545} \approx 0.949$$

Kesesuaian dengan D_2 (Selat Solo Mbak Lies):

$$\vec{V}_Q \cdot \vec{V}_{D2} = (1.000 \times 1.000) = 1.000$$

$$\text{Sim}(Q, D_2) = \frac{1.000}{3.600 \times 1.888} \approx 0.147$$

Kemiripan dengan D_3 (Tahok Pak Citro):

$$\vec{V}_Q \cdot \vec{V}_{D3} = (1.301 \times 1.301) + (1.000 \times 1.000) = 2.693$$

$$\text{Sim}(Q, D_3) = \frac{2.693}{3.600 \times 1.641} \approx 0.456$$

Kemiripan dengan D_4 (Sate Buntel Badran):

$$\vec{V}_Q \cdot \vec{V}_{D4} = (1.000 \times 1.000) = 1.000$$

$$\text{Sim}(Q, D_4) = \frac{1.000}{3.600 \times 1.000} \approx 0.278$$

Nilai kemiripan diurutkan secara menurun untuk memposisikan kuliner yang paling relevan di peringkat teratas. Hasil pemeringkatan rekomendasi disajikan pada Tabel VI.

Tabel 6 Peringkat Hasil Rekomendasi Akhir UMKM Kuliner

Rank	ID	Nama UMKM	Cosine Sim	Status
1	D1	Soto Arinda	0.949	Sangat Relevan
2	D2	Tahok Pak Citro	0.456	Cukup Relevan

3	D3	Sate Buntel Badran	0.278	Kurang Relevan
4	D4	Selat Solo Mbak Lies	0.147	Tidak Relevan

F. Pembahasan

Berdasarkan hasil peringkat pada Tabel VI, **Soto Arinda (Kerten)** (D1) terpilih sebagai rekomendasi utama dengan skor kemiripan 0.949. Skor yang tinggi ini mengindikasikan bahwa sistem mampu mengidentifikasi kesamaan semantik fitur menu ("soto", "sapi", "kuah bening segar") sekaligus mencocokkan ketersediaan fasilitas fisik ("parkir") yang diinputkan secara presisi oleh pengguna. Sebaliknya, **Tahok Pak Citro** (D3) menempati peringkat kedua (0.456) karena adanya kecocokan pada variabel "kuah" dan penyediaan fasilitas "parkir", yang mengonfirmasi ketahanan logika semantik CBF dalam menyaring item relevan.

Penerapan sistem rekomendasi berbasis konten multi-atribut ini memberikan kontribusi yang signifikan bagi perkembangan UMKM kuliner tradisional di Kota Surakarta. Selama ini, banyak UMKM kuliner tradisional berskala mikro memiliki keterbatasan dalam melakukan promosi secara digital sehingga sulit bersaing dengan restoran modern berskala besar. Dengan mengintegrasikan sistem pencocokan berbasis fasilitas tempat makan dan kecocokan rasa, sistem ini mampu memberikan rekomendasi secara objektif. Wisatawan dapat menemukan tempat makan otentik yang memiliki fasilitas pendukung yang mereka butuhkan (seperti tersedianya area parkir atau akses toilet) secara langsung dari basis data Google Maps.

Secara tidak langsung, sistem ini juga mendukung pelestarian budaya kuliner khas Kota Surakarta. Wisatawan tidak hanya memperoleh referensi tempat makan populer, melainkan juga diarahkan pada variasi kuliner tradisional Solo yang beraneka ragam melalui pencocokan semantik kata kunci menu. Digitalisasi fitur konten dan fasilitas fisik UMKM ini membantu meningkatkan ketahanan ekonomi lokal pelaku usaha kuliner di Kota Surakarta.

IV. KESIMPULAN

Penelitian ini berhasil merancang dan mengimplementasikan sistem rekomendasi kuliner berbasis *Content-Based Filtering* (CBF) untuk wilayah Kota Surakarta menggunakan data ulasan objektif dari Google Maps. Atribut fasilitas fisik tempat makan (seperti parkir, toilet, mushola, WiFi, dll.) yang dipadukan dengan fitur kategori dan menu terbukti andal dalam menghasilkan keputusan rekomendasi yang sesuai dengan preferensi kenyamanan dan selera pengguna.

Melalui pemodelan empat tahap yang meliputi akuisisi data, pra-pemrosesan teks, pembobotan TF-IDF, dan pengukuran kemiripan *Cosine Similarity*, sistem mampu memberikan rekomendasi yang presisi dan bebas dari bias penilaian rating subjektif. Hasil komputasi simulasi menunjukkan bahwa sistem berhasil merekomendasikan "Soto Arinda (Kerten)" sebagai kuliner paling relevan (skor 0.949) bagi pengguna yang menginginkan hidangan soto sapi kuah bening segar dengan fasilitas area parkir yang memadai. Sistem rekomendasi ini dapat mempermudah wisatawan dalam menemukan kuliner legendaris di Kota Surakarta sekaligus memperluas jangkauan promosi digital bagi pelaku UMKM kuliner lokal secara adil dan berkelanjutan.

Sebagai saran untuk pengembangan penelitian selanjutnya, sistem rekomendasi ini dapat ditingkatkan dengan menerapkan metrik evaluasi standar recommender system secara komprehensif, seperti pengujian nilai Precision, Recall, Mean Average Precision (MAP), Normalized Discounted Cumulative Gain (NDCG), atau melalui evaluasi langsung terhadap pengguna (user acceptance testing). Jumlah dokumen dalam korpus komputasi pembobotan TF-IDF dan *Cosine Similarity* juga perlu diperluas ke seluruh dataset ($N = 50$) agar keputusan rekomendasi yang dihasilkan lebih merepresentasikan kondisi riil di lapangan. Terakhir, guna meminimalkan bias sumber data yang timbul dari variabilitas kelengkapan informasi mandiri oleh pemilik usaha maupun pengulas di platform Google Maps, penelitian

berikutnya dapat mengombinasikan metode akuisisi data ini dengan survei validasi lapangan langsung serta teknik pembobotan berbasis korelasi multi-kriteria.

V. UCAPAN TERIMA KASIH

Peneliti menyampaikan apresiasi yang tulus kepada Fakultas Ilmu Komputer, Universitas Duta Bangsa Surakarta, atas bantuan sarana akademis yang diberikan, serta kepada komunitas pengulas digital Google Maps Surakarta yang ulasan otentiknya telah membantu mewujudkan basis data kurasi yang andal demi kemajuan ekonomi kreatif daerah.

REFERENSI

- [1] Muchalim Danu Warta, Pramono, and Joni Maulindar, "SISTEM REKOMENDASI KULINER IKONIK KOTA SOLO MENGGUNAKAN METODE CONTENT BASED FILTERING," *J. Teknol. Terpadu*, vol. 11, pp. 78–86, Jul. 2025, doi: <https://doi.org/10.54914/jtt.v11i1.1870>.
- [2] R. H. Mondy and A. Wijayanto, "RECOMMENDATION SYSTEM WITH CONTENT-BASED FILTERING METHOD FOR CULINARY TOURISM IN MANGAN APPLICATION."
- [3] I. Rasyid, M. R. A. Yudianto, M. Maimunah, and T. A. Purnomo, "Electronic Product Recommendation System Using the Cosine Similarity Algorithm and VGG-16," *sinkron*, vol. 8, no. 4, pp. 2120–2129, Oct. 2023, doi: [10.33395/sinkron.v8i4.12936](https://doi.org/10.33395/sinkron.v8i4.12936).
- [4] Z. D. Artika and E. B. Setiawan, "CONTENT-BASED FILTERING CULINARY RECOMMENDATION SYSTEM USING DEEP CONVOLUTIONAL NEURAL NETWORK ON TWITTER (X)," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 2, pp. 333–341, Nov. 2024, doi: [10.33480/jitk.v10i2.5640](https://doi.org/10.33480/jitk.v10i2.5640).
- [5] Z. A. Ekovich, Moh. Muhtarom, and N. F. Muhammad, "Solo Culinary Recommendation System Use Web-Based Collaborative Filtering Method," *J. Adv. Inf. Ind. Technol.*, vol. 6, no. 2, pp. 103–112, Oct. 2024, doi: [10.52435/jait.v6i2.578](https://doi.org/10.52435/jait.v6i2.578).
- [6] N. Saputri and P. Suryati, "Machine Learning-Based Book and Library Recommendation Application Using Content-Based Filtering Method," *J. Teknol. Inf. dan Terap.*, vol. 12, no. 1, pp. 25–33, Jun. 2025, doi: [10.25047/jtit.v12i1.445](https://doi.org/10.25047/jtit.v12i1.445).
- [7] N. P. T. Widiyanti, I. M. A. D. Suarjaya, and N. K. D. Rusjanyanti, "Food Recipe Recommendation System with Content-Based Filtering and Collaborative Filtering Methods," *Sinkron*, vol. 9, no. 3, pp. 1167–1776, Aug. 2025, doi: [10.33395/sinkron.v9i3.14778](https://doi.org/10.33395/sinkron.v9i3.14778).
- [8] Z. F. Falah and F. Suryawan, "Recommendation System to Propose Final Project Supervisor using Cosine Similarity Matrix," 2022.
- [9] R. G. Nugroho, S. Sumarlinda, and A. Srirahayu, "Camping Equipment Recommendation System Using Content-Based Filtering Method: A Case Study of Berkah Outdoor45," *Int. J. Softw. Eng. Comput. Sci.*, vol. 4, no. 3, pp. 1077–1088, Dec. 2024, doi: [10.35870/ijsecs.v4i3.3078](https://doi.org/10.35870/ijsecs.v4i3.3078).
- [10] R. Amelia, A. L. Nurlaili, and F. P. Aditiawan, "A Web-Based Online Reservation System with Personalized Tourism Recommendations Using Content-Based Filtering," *bit-Tech*, vol. 8, no. 3, pp. 3173–3183, Apr. 2026, doi: [10.32877/bt.v8i3.3382](https://doi.org/10.32877/bt.v8i3.3382).

- [11] F. Christyawan, A. N. Rohman, and A. D. Hartanto, "Application of Content-Based Filtering Method Using Cosine Similarity in Restaurant Selection Recommendation System," *J. Inf. Syst. Informatics*, vol. 6, no. 3, pp. 1559–1576, Sep. 2024, doi: 10.51519/journalisi.v6i3.806.