

Sistem Rekomendasi Oleh-Oleh Kerajinan Khas Yogyakarta Menggunakan Pendekatan Content-Based Filtering

**Muhammad Anugrah Putra¹, Muhammad Hanif Hilmi², Reza Aditya Angga Putra³, Ahmad
Bagus Prakoso⁴, Rio Sandy Laksono⁵, Vihi Atina⁶**

¹⁻⁶Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Duta Bangsa, Surakarta,
Indonesia

Email: ¹230103142@mhs.udb.ac.id, ²230103143@mhs.udb.ac.id, ³230103149@mhs.udb.ac.id,
⁴230103124@mhs.udb.ac.id, ⁵230103269@mhs.udb.ac.id, ⁶vihi_atina@udb.ac.id

Abstrak—Keberagaman produk oleh-oleh kerajinan khas Yogyakarta menyebabkan wisatawan dan masyarakat mengalami kesulitan dalam menemukan produk yang sesuai dengan kebutuhan dan preferensi karena proses pencarian masih dilakukan secara manual. Penelitian ini bertujuan merancang dan mengimplementasikan sistem rekomendasi oleh-oleh kerajinan khas Yogyakarta menggunakan pendekatan Content-Based Filtering untuk membantu pengguna memperoleh produk yang paling relevan berdasarkan karakteristiknya. Metode pelaksanaan dilakukan melalui pengumpulan dataset produk, preprocessing data, pembobotan menggunakan Term Frequency–Inverse Document Frequency (TF-IDF), perhitungan tingkat kemiripan dengan Cosine Similarity, serta implementasi sistem menggunakan framework Laravel 9 dan basis data MySQL. Hasil penelitian menunjukkan bahwa sistem berhasil mengolah atribut produk berupa kategori, bahan, lokasi asal, dan fungsi produk sebagai representasi karakteristik setiap produk. Proses preprocessing mampu menghasilkan data yang lebih terstruktur sehingga mendukung pembentukan vektor TF-IDF secara optimal. Perhitungan Cosine Similarity berhasil menentukan tingkat kemiripan antarproduk berdasarkan nilai similarity yang dihasilkan. Sistem mampu menampilkan lima rekomendasi produk dengan tingkat kemiripan tertinggi terhadap produk yang dipilih pengguna, seperti pada produk kerajinan batik “Batik Kawung Elegan” yang merekomendasikan produk “Batik Tulis Giriloyo” dengan nilai similarity tertinggi sebesar 0,4453 dan produk “Syal Batik Motif Sekar Jagad” dengan nilai similarity 0,3316. Implementasi antarmuka yang terdiri atas halaman katalog, halaman detail produk beserta rekomendasi, dan halaman administrasi berjalan sesuai dengan kebutuhan fungsional sistem. Penyimpanan matriks kemiripan pada basis data membuat proses penyajian rekomendasi menjadi lebih cepat karena tidak memerlukan perhitungan ulang setiap kali pengguna mengakses aplikasi. Secara keseluruhan, sistem rekomendasi yang dikembangkan mampu membantu pengguna menemukan alternatif oleh-oleh kerajinan khas Yogyakarta secara lebih efektif sekaligus mendukung promosi produk UMKM melalui penyajian rekomendasi berbasis kesamaan karakteristik produk.

Kata Kunci: *Content-Based Filtering, Cosine Similarity, Laravel, Sistem Rekomendasi, TF-IDF*

Abstract—The wide variety of Yogyakarta handicraft souvenirs makes it difficult for tourists and the public to find products that match their preferences because the selection process is still carried out manually. This study aims to design and implement a recommendation system for Yogyakarta handicraft souvenirs using the Content-Based Filtering approach to assist users in finding products that best match their preferences. The implementation method involved collecting product datasets, performing data preprocessing, applying Term Frequency–Inverse Document Frequency (TF-IDF) weighting, calculating product similarity using Cosine Similarity, and developing the system with the Laravel 9 framework and MySQL database. The results indicate that the system successfully processes product attributes, including category, material, origin, and product function, as representations of product characteristics. The preprocessing stage effectively organizes the data to support the generation of accurate TF-IDF vectors. The Cosine Similarity method successfully measures the similarity between products based on their attribute vectors. The system is able to provide the top five product recommendations with the highest similarity scores for each selected product. The implemented user interface, consisting of the product catalog, product detail page with recommendations, and administrator page, operates according to the defined functional requirements. Storing the similarity matrix in the database improves recommendation efficiency by eliminating the need for repeated calculations whenever users access the system. Overall, the developed recommendation system effectively assists users in discovering relevant Yogyakarta handicraft souvenirs while supporting the promotion of local MSME products through recommendations based on product characteristics.

Keywords: *Content-Based Filtering, Cosine Similarity, Laravel, Recommendation System, TF-IDF.*

1. PENDAHULUAN

Daerah Istimewa Yogyakarta merupakan salah satu tujuan wisata terpopuler di Indonesia yang dikenal luas karena kekayaan budaya, sejarah, seni, dan tradisi yang masih dilestarikan hingga

saat ini(Alamsyah et al., 2023). Selain memiliki berbagai destinasi wisata yang menarik, Yogyakarta juga menjadi sentra industri kreatif yang menghasilkan beragam produk kerajinan khas, seperti batik, gerabah Kasongan, kerajinan perak Kotagede, ukiran kayu, anyaman bambu, serta berbagai jenis suvenir yang banyak dipilih wisatawan sebagai oleh-oleh. Meningkatnya aktivitas sektor pariwisata turut memberikan dampak positif terhadap perkembangan Usaha Mikro, Kecil, dan Menengah (UMKM) di bidang kerajinan(Rivalni et al., 2025). Kondisi tersebut mendorong munculnya berbagai produk dengan karakteristik yang beragam, baik ditinjau dari kategori, bahan baku, kisaran harga, daerah asal, maupun fungsi produk. Keberagaman ini memang memperluas pilihan bagi konsumen, namun sekaligus menimbulkan kesulitan dalam menentukan produk yang paling sesuai dengan kebutuhan dan preferensi masing-masing.

Dalam praktiknya, wisatawan maupun masyarakat yang hendak membeli oleh-oleh kerajinan umumnya masih memperoleh informasi melalui rekomendasi kerabat, media sosial, ataupun melakukan pencarian secara manual pada berbagai platform digital. Cara tersebut dinilai kurang efisien karena calon pembeli harus menelusuri dan membandingkan banyak produk sebelum menentukan pilihan. Tidak sedikit konsumen yang akhirnya memutuskan pembelian hanya berdasarkan tingkat popularitas atau harga produk tanpa mempertimbangkan kesesuaian karakteristik yang sebenarnya diinginkan. Seiring bertambahnya jumlah produk kerajinan yang dipasarkan, baik secara langsung maupun melalui platform digital, proses pencarian menjadi semakin kompleks sehingga diperlukan cara yang lebih efektif untuk membantu pengguna menemukan produk yang relevan.

Permasalahan tersebut juga dirasakan oleh para pelaku UMKM kerajinan. Sebagian besar produk yang dimiliki sering kali tidak seluruhnya diketahui oleh calon pembeli karena hanya beberapa produk tertentu yang ditampilkan pada halaman utama toko atau marketplace(Rahmawati & Suwarni, 2023). Akibatnya, produk lain yang memiliki karakteristik serupa dan berpotensi menarik perhatian konsumen tidak memperoleh kesempatan yang sama untuk dipromosikan. Kondisi ini menyebabkan peluang penjualan belum dapat dimaksimalkan. Oleh karena itu, diperlukan suatu sistem yang mampu menyajikan rekomendasi produk berdasarkan kesamaan karakteristik sehingga konsumen lebih mudah menemukan alternatif produk yang sesuai, sekaligus membantu pelaku usaha meningkatkan eksposur terhadap koleksi produk yang mereka miliki.

Salah satu teknologi yang dapat diterapkan untuk mengatasi permasalahan tersebut adalah sistem rekomendasi. Sistem rekomendasi berfungsi memberikan saran produk kepada pengguna berdasarkan informasi atau karakteristik tertentu sehingga proses pemilihan menjadi lebih cepat, efektif, dan sesuai dengan kebutuhan pengguna. Dalam berbagai aplikasi perdagangan elektronik, sistem rekomendasi telah banyak dimanfaatkan untuk meningkatkan pengalaman pengguna, mempermudah proses pencarian produk, serta meningkatkan peluang terjadinya transaksi. Penerapan sistem rekomendasi pada produk kerajinan khas Yogyakarta diharapkan mampu membantu wisatawan maupun masyarakat memperoleh alternatif oleh-oleh yang sesuai dengan preferensi mereka tanpa harus melakukan pencarian secara manual terhadap seluruh produk yang tersedia.

Penelitian ini menerapkan pendekatan Content-Based Filtering (CBF) sebagai metode dalam membangun sistem rekomendasi(Rachmaniar et al., 2025a)(Pranata & Sulianta, 2025). Pendekatan tersebut menghasilkan rekomendasi dengan memanfaatkan karakteristik yang dimiliki setiap produk, kemudian mencari produk lain yang mempunyai tingkat kemiripan tinggi berdasarkan atribut-atribut tersebut. Atribut yang digunakan dalam penelitian meliputi kategori produk, bahan utama, lokasi asal, dan fungsi produk. Seluruh atribut tersebut digabungkan menjadi representasi konten yang selanjutnya diproses menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk menghasilkan bobot setiap fitur. Setelah proses pembobotan selesai, tingkat kemiripan antarproduk dihitung menggunakan metode Cosine Similarity, sehingga sistem mampu menyajikan rekomendasi produk yang memiliki karakteristik paling mendekati produk yang dipilih oleh pengguna.

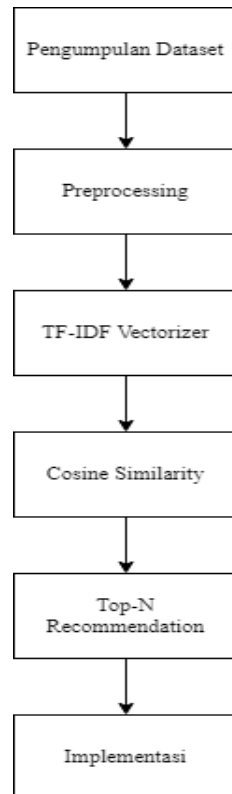
Pemanfaatan pendekatan Content-Based Filtering pada penelitian ini didasarkan pada kemampuannya dalam menghasilkan rekomendasi tanpa bergantung pada data riwayat transaksi maupun penilaian pengguna. Karakteristik tersebut sesuai dengan dataset yang digunakan, yaitu data produk kerajinan khas Yogyakarta yang memuat informasi mengenai nama produk, kategori, bahan, harga, lokasi asal, dan fungsi produk. Dengan memanfaatkan atribut-atribut tersebut sebagai dasar

pembentukan profil produk, sistem diharapkan mampu menghasilkan rekomendasi yang lebih relevan dan akurat. Selain memberikan kemudahan bagi konsumen dalam memilih oleh-oleh yang sesuai dengan preferensi mereka, sistem yang dikembangkan juga diharapkan dapat menjadi media pendukung promosi produk kerajinan lokal serta membantu meningkatkan daya saing UMKM Yogyakarta melalui penerapan teknologi sistem rekomendasi berbasis konten.

Penelitian ini bertujuan untuk merancang dan mengembangkan sistem rekomendasi oleh-oleh kerajinan khas Yogyakarta menggunakan pendekatan Content-Based Filtering sebagai solusi untuk membantu pengguna memperoleh produk yang sesuai dengan preferensi dan karakteristik yang diinginkan (Praditya et al., 2021). Sistem dibangun dengan memanfaatkan atribut produk yang meliputi kategori, bahan, lokasi asal, dan fungsi produk sebagai representasi konten dari setiap produk kerajinan. Selanjutnya, metode TF-IDF digunakan untuk memberikan bobot terhadap setiap atribut sehingga karakteristik produk dapat direpresentasikan dalam bentuk vektor numerik (Anggraeni et al., 2026). Tingkat kemiripan antarproduk kemudian dihitung menggunakan Cosine Similarity untuk menghasilkan daftar rekomendasi produk yang memiliki tingkat relevansi tertinggi. Selain itu, penelitian ini juga bertujuan mengevaluasi kemampuan pendekatan Content-Based Filtering dalam menghasilkan rekomendasi berdasarkan kesamaan atribut produk tanpa memerlukan data histori transaksi maupun penilaian pengguna. Hasil yang diperoleh diharapkan dapat membantu wisatawan dan masyarakat memilih oleh-oleh kerajinan khas Yogyakarta secara lebih cepat dan tepat, sekaligus menjadi sarana pendukung promosi bagi pelaku UMKM agar produk yang mereka miliki lebih mudah ditemukan oleh calon konsumen. Dengan demikian, sistem rekomendasi yang dikembangkan diharapkan mampu meningkatkan kualitas layanan pencarian produk, mempermudah proses pengambilan keputusan pembelian, serta mendukung digitalisasi pemasaran produk kerajinan lokal di Daerah Istimewa Yogyakarta.

2. METODE PENELITIAN

Penelitian ini mengembangkan sistem rekomendasi oleh-oleh kerajinan khas Yogyakarta menggunakan pendekatan Content-Based Filtering (CBF). Pendekatan ini dipilih karena mampu menghasilkan rekomendasi berdasarkan karakteristik setiap produk tanpa memerlukan riwayat transaksi maupun penilaian pengguna (Rachmaniar et al., 2025b). Dataset yang digunakan terdiri atas data kerajinan khas Yogyakarta yang memiliki atribut nama produk, kategori, bahan, lokasi asal, fungsi produk, dan harga. Pada penelitian ini atribut nama produk, kategori, bahan, lokasi, dan fungsi digunakan sebagai representasi karakteristik produk, sedangkan atribut harga hanya digunakan sebagai informasi pendukung pada hasil rekomendasi. Proses penelitian dimulai dari pengumpulan dataset, dilanjutkan dengan tahap preprocessing, pembentukan bobot menggunakan TF-IDF Vectorizer, perhitungan tingkat kemiripan menggunakan Cosine Similarity, hingga menghasilkan rekomendasi produk berdasarkan nilai kemiripan tertinggi. Tahapan penelitian ditunjukkan pada Gambar 1.



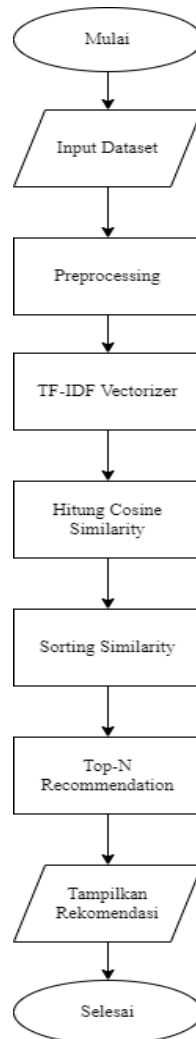
Gambar 1. Tahapan Penelitian

Tahapan penelitian dimulai dengan pengumpulan dataset kerajinan khas Yogyakarta yang diperoleh dari berbagai sumber. Selanjutnya dilakukan preprocessing untuk membersihkan data teks sehingga menghasilkan data yang lebih terstruktur. Data yang telah diproses kemudian diubah menjadi representasi numerik menggunakan TF-IDF Vectorizer. Vektor TF-IDF tersebut digunakan untuk menghitung tingkat kemiripan antarproduk menggunakan Cosine Similarity. Berdasarkan nilai kemiripan yang diperoleh, sistem melakukan pengurutan (*sorting*) sehingga diperoleh sejumlah produk dengan nilai similarity tertinggi yang ditampilkan sebagai rekomendasi kepada pengguna.

2.1 Flowchart Sistem

Flowchart sistem digunakan untuk menggambarkan proses kerja sistem rekomendasi secara keseluruhan mulai dari pengguna memilih produk hingga sistem menghasilkan rekomendasi. Flowchart memberikan gambaran mengenai urutan proses yang dilakukan oleh sistem sehingga memudahkan dalam memahami mekanisme rekomendasi yang diterapkan pada penelitian ini.

Proses diawali ketika pengguna memilih salah satu produk kerajinan sebagai produk acuan (*input product*). Selanjutnya sistem membaca seluruh dataset dan melakukan preprocessing terhadap atribut teks yang digunakan sebagai representasi produk. Setelah data berhasil diproses, sistem membentuk representasi numerik menggunakan TF-IDF Vectorizer. Representasi tersebut kemudian digunakan untuk menghitung tingkat kemiripan antara produk yang dipilih pengguna dengan seluruh produk yang terdapat pada dataset menggunakan metode Cosine Similarity. Nilai kemiripan yang diperoleh kemudian diurutkan dari nilai terbesar hingga terkecil menggunakan proses *sorting*. Produk dengan nilai similarity tertinggi dipilih sebagai rekomendasi karena dianggap memiliki karakteristik yang paling mendekati produk yang dipilih pengguna. Hasil rekomendasi kemudian ditampilkan dalam bentuk daftar produk yang dapat dijadikan alternatif oleh pengguna.



Gambar 2. Flowchart Sistem

Berdasarkan flowchart tersebut dapat diketahui bahwa seluruh proses dilakukan secara berurutan mulai dari pengolahan data hingga menghasilkan rekomendasi. Setiap tahapan saling berkaitan sehingga kualitas hasil rekomendasi sangat dipengaruhi oleh keberhasilan proses preprocessing, pembobotan TF-IDF, dan perhitungan Cosine Similarity.

2.2 Preprocessing Data

Tahap preprocessing merupakan tahap awal yang sangat penting dalam penelitian ini karena bertujuan menghasilkan data yang bersih, seragam, dan siap digunakan pada proses pembobotan. Data produk yang diperoleh dari berbagai sumber umumnya masih mengandung karakter yang tidak diperlukan, seperti huruf kapital, tanda baca, angka, simbol tertentu, maupun kata-kata yang tidak memiliki makna penting. Oleh karena itu diperlukan proses pembersihan data agar informasi yang digunakan benar-benar merepresentasikan karakteristik produk.

Pada penelitian ini atribut nama produk, kategori, bahan, lokasi asal, dan fungsi produk digabungkan menjadi satu dokumen yang merepresentasikan setiap produk kerajinan. Penggabungan atribut dilakukan agar seluruh karakteristik produk dapat diproses secara bersamaan sehingga menghasilkan representasi yang lebih lengkap dibandingkan apabila setiap atribut diproses secara terpisah.

Tahap preprocessing terdiri atas beberapa proses, yaitu case folding, cleaning, tokenizing, dan stopword removal. Case folding dilakukan dengan mengubah seluruh huruf menjadi huruf

kecil (*lowercase*) sehingga tidak terjadi perbedaan antara huruf kapital dan huruf kecil. Tahap cleaning dilakukan untuk menghapus angka, karakter khusus, tanda baca, serta spasi berlebih yang tidak memiliki kontribusi terhadap proses rekomendasi. Selanjutnya dilakukan tokenizing untuk memecah kalimat menjadi kumpulan kata (*terms*). Tahap terakhir adalah stopword removal yang bertujuan menghilangkan kata-kata umum seperti *dan*, *yang*, *di*, *ke*, *dari*, dan kata lain yang tidak memberikan informasi penting terhadap karakteristik produk. Hasil preprocessing berupa sekumpulan kata penting yang memiliki makna representatif terhadap setiap produk. Kumpulan kata tersebut selanjutnya digunakan sebagai masukan pada proses pembobotan TF-IDF.

2.3 Content-Based Filtering

Content-Based Filtering merupakan metode sistem rekomendasi yang bekerja dengan membandingkan karakteristik setiap item berdasarkan atribut yang dimiliki. Berbeda dengan Collaborative Filtering yang membutuhkan data histori pengguna, Content-Based Filtering hanya memanfaatkan informasi yang terdapat pada produk sehingga sangat sesuai diterapkan pada penelitian ini (Phalle & Bhushan, 2024).

Dalam penelitian ini setiap produk direpresentasikan sebagai sebuah dokumen yang terdiri atas atribut nama produk, kategori, bahan, lokasi asal, dan fungsi produk. Sistem akan membandingkan dokumen produk yang dipilih pengguna dengan seluruh dokumen pada dataset untuk mengetahui tingkat kemiripan karakteristik antarproduk.

Apabila dua produk memiliki kata-kata penting yang sama atau memiliki karakteristik yang serupa, maka kedua produk tersebut akan memperoleh nilai similarity yang tinggi. Sebaliknya apabila karakteristik kedua produk berbeda, maka nilai similarity yang dihasilkan akan semakin kecil. Dengan demikian sistem dapat memberikan rekomendasi yang sesuai dengan karakteristik produk yang sedang dicari oleh pengguna.

Keunggulan pendekatan ini adalah tidak memerlukan data transaksi maupun penilaian pengguna sehingga mampu mengatasi permasalahan cold-start user. Selain itu rekomendasi yang dihasilkan lebih mudah dijelaskan karena didasarkan pada atribut yang benar-benar dimiliki oleh setiap produk.

2.4 TF-IDF Vectorizer

Setelah proses preprocessing selesai dilakukan, setiap dokumen diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot terhadap setiap kata berdasarkan tingkat kepentingannya pada suatu dokumen.

Tahap pertama adalah menghitung Term Frequency (TF) yang menunjukkan seberapa sering suatu kata muncul pada sebuah dokumen. Semakin sering suatu kata muncul, maka nilai TF yang dihasilkan akan semakin besar.

Persamaan (1)

$$TF(t, d) = \frac{f(t, d)}{\sum f(d)}$$

Tahap berikutnya adalah menghitung Inverse Document Frequency (IDF) untuk mengetahui tingkat keunikan suatu kata pada seluruh dokumen. Kata yang muncul pada banyak dokumen akan memperoleh bobot lebih kecil dibandingkan kata yang hanya muncul pada beberapa dokumen.

Persamaan (2)

$$IDF(t) = \log \left(\frac{N}{df(t)} \right)$$

Bobot akhir TF-IDF diperoleh dengan mengalikan nilai TF dan IDF sebagaimana ditunjukkan pada Persamaan (3).

$$TFIDF(t, d) = TF(t, d) \times IDF(t)$$

Hasil pembobotan TF-IDF berupa vektor numerik yang merepresentasikan setiap produk berdasarkan karakteristik kata-kata penting yang dimilikinya. Vektor inilah yang digunakan sebagai dasar dalam proses penghitungan tingkat kemiripan antarproduk menggunakan metode Cosine Similarity.

2.5 Cosine Similarity

Setelah seluruh produk direpresentasikan dalam bentuk vektor TF-IDF, tahap berikutnya adalah menghitung tingkat kemiripan menggunakan metode Cosine Similarity. Metode ini mengukur sudut antara dua buah vektor sehingga menghasilkan nilai kemiripan dalam rentang 0 hingga 1. Nilai yang mendekati 1 menunjukkan bahwa kedua produk memiliki karakteristik yang hampir sama, sedangkan nilai yang mendekati 0 menunjukkan bahwa kedua produk memiliki karakteristik yang berbeda. Oleh karena itu metode Cosine Similarity banyak digunakan pada sistem rekomendasi berbasis teks karena mampu memberikan hasil yang cukup akurat.

Persamaan (4)

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

Nilai kemiripan dihitung terhadap seluruh pasangan produk sehingga diperoleh sebuah matriks similarity. Matriks tersebut menjadi dasar dalam menentukan produk mana yang memiliki karakteristik paling mirip dengan produk yang dipilih pengguna.

2.6 Perhitungan Rekomendasi

Tahap terakhir pada penelitian ini adalah menghasilkan rekomendasi berdasarkan nilai Cosine Similarity yang telah diperoleh. Seluruh nilai similarity diurutkan dari nilai terbesar hingga terkecil menggunakan proses *sorting* (Timur, 2025). Pengurutan dilakukan untuk memperoleh produk yang memiliki tingkat kemiripan paling tinggi terhadap produk acuan.

Penelitian ini menerapkan konsep Top-N Recommendation, yaitu hanya menampilkan sejumlah produk terbaik kepada pengguna. Pada implementasinya digunakan Top-5 Recommendation, sehingga sistem akan menampilkan lima produk dengan nilai similarity tertinggi sebagai hasil rekomendasi.

Informasi yang ditampilkan kepada pengguna meliputi nama produk, kategori, bahan, lokasi asal, fungsi produk, harga, serta nilai similarity masing-masing produk. Dengan adanya informasi tersebut pengguna tidak hanya memperoleh rekomendasi produk yang relevan, tetapi juga dapat membandingkan karakteristik antarproduk sebelum menentukan pilihan. Pendekatan ini diharapkan mampu membantu wisatawan maupun masyarakat dalam menemukan oleh-oleh kerajinan khas Yogyakarta yang sesuai dengan kebutuhan dan preferensi mereka.

3. ANALISA DAN PEMBAHASAN

Pada bab ini dibahas hasil implementasi sistem rekomendasi oleh-oleh kerajinan khas Yogyakarta menggunakan pendekatan Content-Based Filtering (CBF) (Siahaan et al., 2024). Analisis dilakukan terhadap proses pembangunan sistem, pengolahan data, perhitungan TF-IDF, perhitungan Cosine Similarity, serta implementasi antarmuka pengguna berbasis Laravel 9.

3.1 Pengumpulan Dataset

Tahap pertama penelitian ini adalah pengumpulan dataset produk oleh-oleh kerajinan khas Yogyakarta yang diperoleh dari berbagai sumber. Dataset yang terkumpul terdiri atas produk kerajinan dari lima kategori utama, yaitu Batik, Gerabah, Perhiasan, Kerajinan Kayu, dan Anyaman.

Setiap produk pada dataset memiliki atribut nama produk, kategori, bahan, harga, lokasi asal, dan fungsi produk. Contoh data produk kerajinan yang telah dikumpulkan ditunjukkan pada Tabel 1.

Tabel 1. Dataset Kerajinan

Nama produk	Kategori	Bahan	Harga (Rp)	Lokasi	Fungsi produk	
Batik Parang Premium	Batik	Katun	350000	Bantul	Fashion	
Batik Kawung Elegan	Batik	Katun	275000	Bantul	Fashion	
Batik Mega Mendung Jogja	Batik	Sutra	650000	Yogyakarta	Fashion	
Batik Tulis Giriloyo	Batik	Katun	450000	Bantul	Fashion	
Syal Motif Sekar Jagad	Batik	Katun	180000	Sleman	Fashion	
Vas Gerabah Kasongan	Gerabah	Tanah Liat	180000	Bantul	Dekorasi	
Pot Bunga Gerabah	Gerabah	Tanah Liat	120000	Bantul	Dekorasi	
Celengan Gerabah	Gerabah	Tanah Liat	85000	Bantul	Souvenir	
Tempat Lilin Gerabah	Gerabah	Tanah Liat	95000	Bantul	Dekorasi	
Teko Gerabah Tradisional	Gerabah	Tanah Liat	210000	Bantul	Peralatan Rumah Tangga	
Kalung Perak Kotagede	Perhiasan	Perak	450000	Kotagede	Aksesoris	
Gelang Perak Motif Jawa	Perhiasan	Perak	320000	Kotagede	Aksesoris	
Anting Perak Handmade	Perhiasan	Perak	275000	Kotagede	Aksesoris	
Bros Bunga	Perak	Perak	190000	kotagede	Aksesoris	
Cincin Etnik	Perak	Perak	380000	Kotagede	Aksesoris	
Topeng Bobung	Kayu	Kerajinan Kayu	Kayu Mahoni	250000	Gunungkidul	Dekorasi
Miniatur Becak Kayu	Kayu	Kerajinan Kayu	Kayu Jati	150000	Sleman	Souvenir
Ukiran Burung Garuda	Kayu	Kerajinan Kayu	Kayu Jati	520000	Gunungkidul	Dekorasi

Gantungan Kunci Kayu	Kerajinan Kayu	Kayu Pinus	35000	Sleman	Souvenir
Kotak Ukir	Tisu Kerajinan Kayu	Kayu Mahoni	175000	Gunungkidul	Peralatan Rumah Tangga
Tas Anyaman Bambu	Anyaman	Bambu	200000	Kulon Progo	Fashion
Keranjang Anyaman	Anyaman	Bambu	165000	Kulon Progo	Peralatan Rumah Tangga
Tempat Buah Anyaman	Anyaman	Rotan	145000	Kulon Progo	Peralatan Rumah Tangga
Topi Anyaman Pandan	Anyaman	Daun Pandan	110000	Kulon Progo	Fashion
Dompet Anyaman Bambu	Anyaman	Bambu	95000	Kulon Progo	Aksesoris

Data pada Tabel 1 tersebut menjadi masukan utama bagi tahap preprocessing pada bagian berikutnya, di mana atribut kategori, bahan, lokasi, dan fungsi produk akan digabungkan menjadi representasi dokumen teks untuk setiap produk.

3.2 Preprocessing

Tahap preprocessing diimplementasikan pada RecommendationService dan mengikuti tahapan case folding, cleaning, tokenizing, dan stopword removal sebagaimana dijelaskan pada bagian metode penelitian. Setiap produk direpresentasikan sebagai dokumen teks yang dibentuk dari penggabungan atribut kategori, bahan, lokasi, dan fungsi produk, kemudian seluruh teks diubah menjadi huruf kecil (case folding) agar proses perbandingan kata menjadi lebih konsisten. Sebagai contoh, produk dengan kategori Gerabah, bahan Tanah Liat, lokasi Kasongan, dan fungsi Hiasan Rumah akan melalui tahap cleaning untuk menghapus tanda baca dan karakter yang tidak diperlukan, dilanjutkan dengan tokenizing untuk memecah teks menjadi kumpulan kata, serta stopword removal untuk menghilangkan kata-kata umum yang tidak memiliki makna penting. Hasil akhir dari proses tersebut adalah representasi dokumen “gerabah tanah liat kasongan hiasan rumah”.

Representasi tersebut memungkinkan sistem membandingkan karakteristik setiap produk secara langsung berdasarkan kata-kata penting yang dimiliki.

3.3 TF-IDF

Setelah dokumen terbentuk, sistem menghitung Term Frequency (TF) untuk mengetahui seberapa sering suatu kata muncul pada sebuah produk. Selanjutnya dihitung Inverse Document Frequency (IDF) untuk menentukan tingkat keunikan kata pada seluruh dataset. Bobot akhir setiap kata diperoleh menggunakan metode TF-IDF. Kata yang sering muncul pada suatu produk tetapi jarang muncul pada produk lain akan memperoleh bobot yang lebih besar. Dengan demikian, atribut yang benar-benar merepresentasikan karakteristik produk akan memiliki pengaruh lebih tinggi terhadap hasil rekomendasi.

Tabel 2. TF-IDF

Kata	Df	Idf	Kata	Df	Idf
aksesori	6	0.61979	gelang	1	1.39794
anting	1	1.39794	gerabah	5	0.69897
anyaman	5	0.69897	giriloyo	1	1.39794

bambu	3	0.92082	gunungkidul	3	0.92082
bantul	8	0.92082	handmade	1	1.39794
batik	5	0.69897	jagad	1	1.39794
becak	1	1.39794	jati	2	1.09691
bobung	1	1.39794	jawa	1	1.39794
bros	1	1.39794	jogja	1	1.39794
buah	1	1.39794	kalung	1	1.39794
bunga	2	1.09691	kasongan	1	1.39794
burung	1	1.39794	katun	4	0.79588
celengan	1	1.39794	kawung	1	1.39794
cincin	1	1.39794	kayu	5	0.69897
daun	1	1.39794	kerajinan	5	0.69897
dekorasi	5	0.69897	keranjang	1	1.39794
dompet	1	1.39794	kotagede	5	0.69897
elegan	1	1.39794	kotak	1	1.39794
etnik	1	1.39794	kulon	5	0.69897
fashion	7	0.55267	kunci	1	1.39794
garuda	1	1.39794	liat	5	0.69897
gantungan	1	1.39794	lilin	1	1.39794
mahoni	2	1.09691	perhiasan	5	0.69897
mega	1	1.39794	pinus	1	1.39794
mendung	1	1.39794	pot	1	1.39794
miniatur	1	1.39794	premium	1	1.39794
motif	2	1.09691	progo	5	0.69897
pandan	1	1.39794	rotan	1	1.39794
parang	1	1.39794	rumah	4	0.79588
perak	5	0.69897	sekar	1	1.39794
peralatan	4	0.79588	sleman	3	0.92082
souvenir	3	0.92082	tisu	1	1.39794
sutra	1	1.39794	topeng	1	1.39794
syal	1	1.39794	topi	1	1.39794
tanah	5	0.69897	tradisional	1	1.39794
tangga	4	0.79588	tulis	1	1.39794
tas	1	1.39794	ukir	1	1.39794
teko	1	1.39794	ukiran	1	1.39794
tempat	2	1.09691	vas	1	1.39794
yogyakarta	1	1.39794			

Implementasi perhitungan TF-IDF dilakukan secara otomatis pada RecommendationService. Sistem terlebih dahulu menghitung frekuensi kemunculan setiap kata, kemudian menghitung nilai IDF menggunakan logaritma jumlah dokumen dibagi jumlah dokumen yang mengandung kata tersebut. Hasil perkalian TF dan IDF membentuk vektor numerik yang merepresentasikan setiap produk.

3.4 Cosine Similarity

Vektor TF-IDF yang telah dihasilkan kemudian digunakan untuk menghitung tingkat kemiripan antarproduk menggunakan metode Cosine Similarity. Metode ini mengukur sudut antara dua vektor sehingga menghasilkan nilai kemiripan dalam rentang 0 hingga 1.

Tabel 3. Cosine Similarity

ID Produk B	Nama Produk B	Dot Product (A-B)	A	B	Cos θ
2	Batik Kawung Elegan	3,13798	2,6545	2,6545	0,4453

3	Batik Mega Mendung Jogja	2,25968	2,6545	3,4686	0,2454
4	Batik Tulis Giriloyo	3,13798	2,6545	2,6545	0,4453
5	Syal Batik Motif Sekar Jagad	2,89310	2,6545	3,2874	0,3316
6	Vas Gerabah Kasongan	0,24488	2,6545	2,7520	0,0335
7	Pot Bunga Gerabah	0,24488	2,6545	2,6120	0,0353
8	Celengan Gerabah	0,24488	2,6545	2,4451	0,0377
9	Tempat Lilin Gerabah	0,24488	2,6545	2,7520	0,0335
10	Teko Gerabah Tradisional	0,24488	2,6545	2,9975	0,0308
11	Kalung Perak Kotagede	0,00000	2,6545	2,9975	0,0000
12	Gelang Perak Motif Jawa	0,00000	2,6545	2,9030	0,0000
13	Anting Perak Handmade	0,00000	2,6545	2,6877	0,0000
14	Bros Bunga Perak	0,00000	2,6545	2,5440	0,0000
15	Cincin Perak Etnik	0,00000	2,6545	2,6877	0,0000
16	Topeng Kayu Bobung	0,00000	2,6545	3,3666	0,0000
17	Miniatur Becak Kayu	0,00000	2,6545	3,4194	0,0000
18	Ukiran Burung Garuda	0,00000	2,6545	3,2932	0,0000
19	Gantungan Kunci Kayu	0,00000	2,6545	3,5276	0,0000
20	Kotak Tisu Ukir	0,00000	2,6545	3,5010	0,0000
21	Tas Anyaman Bambu	0,30544	2,6545	2,9296	0,0393
22	Keranjang Anyaman	0,00000	2,6545	2,7629	0,0000
23	Tempat Buah Anyaman	0,00000	2,6545	3,1533	0,0000

24	Topi Anyaman Pandan	0,30544	2,6545	3,8681	0,0297
25	Dompot Anyaman Bambu	0,00000	2,6545	2,9430	0,0000

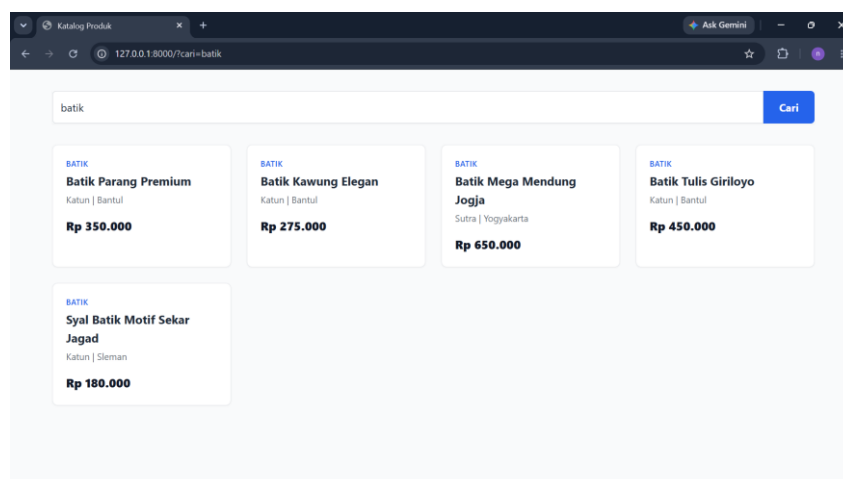
Jika dua produk memiliki banyak atribut yang sama, maka nilai kemiripannya akan mendekati 1. Sebaliknya, jika karakteristik kedua produk berbeda, maka nilai kemiripannya akan mendekati 0. Implementasi algoritma dilakukan melalui fungsi `cosineSimilarity()` yang menghitung dot product dan magnitudo kedua vektor. Hasil perhitungan kemudian disimpan ke tabel `product_similarities` sebagai matriks kemiripan.

3.5 Top N Recommendation

Penelitian ini menerapkan konsep Top-N Recommendation pada tahap implementasi, yaitu hanya menampilkan sejumlah produk dengan nilai similarity tertinggi kepada pengguna. Sistem menggunakan Top-5 Recommendation, sehingga lima produk dengan nilai Cosine Similarity tertinggi ditampilkan sebagai hasil rekomendasi untuk setiap produk yang dipilih pengguna.

Berdasarkan hasil implementasi, sistem mampu menghasilkan rekomendasi produk yang memiliki karakteristik serupa dengan produk yang dipilih pengguna. Sebagai contoh, ketika pengguna memilih produk kerajinan gerabah dari Kasongan, sistem akan merekomendasikan produk lain yang memiliki kategori gerabah, bahan tanah liat, atau berasal dari lokasi yang sama.

Hasil tersebut menunjukkan bahwa pendekatan Content-Based Filtering berhasil memanfaatkan atribut produk untuk menemukan kemiripan antarproduk tanpa memerlukan data histori transaksi maupun penilaian pengguna. Dengan demikian, sistem dapat membantu wisatawan dan masyarakat memperoleh alternatif oleh-oleh kerajinan khas Yogyakarta yang lebih relevan dengan preferensi mereka.



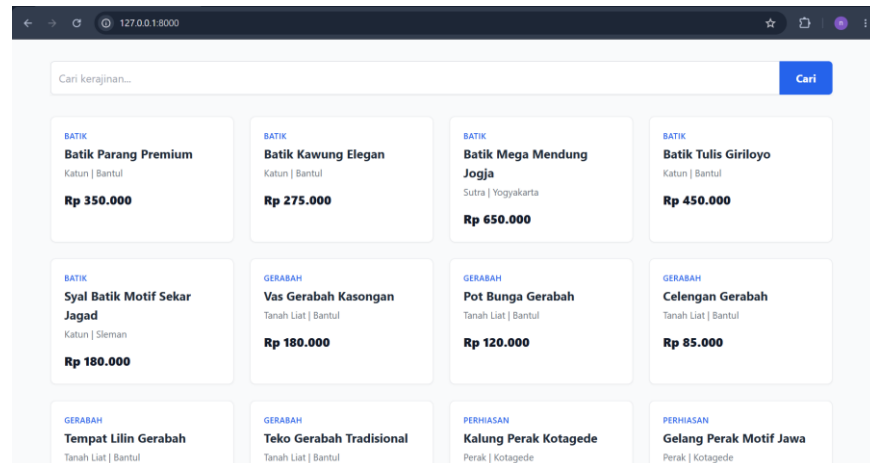
Gambar 3. Rekomendasi

3.6 Analisis Implementasi Antarmuka Sistem

Antarmuka sistem terdiri dari tiga bagian utama, yaitu halaman katalog, halaman detail produk, dan halaman admin.

a. Halaman Katalog Produk

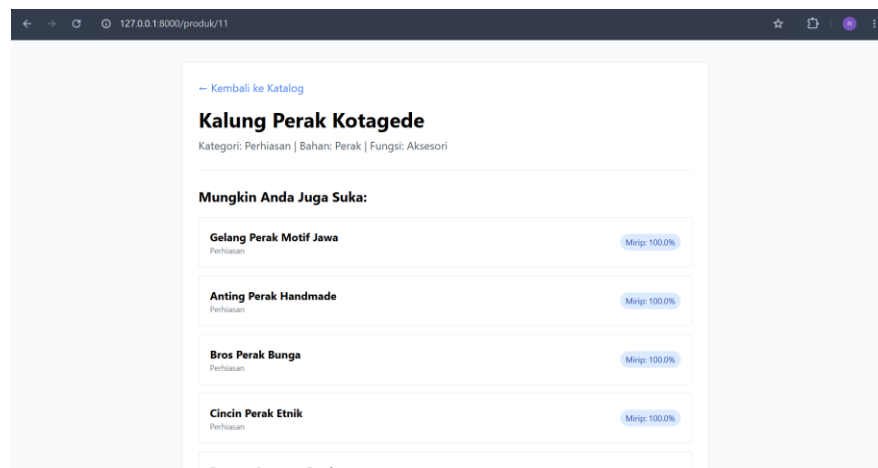
Halaman katalog menampilkan seluruh produk kerajinan beserta fitur pencarian. Pengguna dapat mencari produk berdasarkan nama sehingga proses menemukan produk menjadi lebih cepat.



Gambar 4. Katalog Produk

b. Halaman Detail dan Rekomendasi

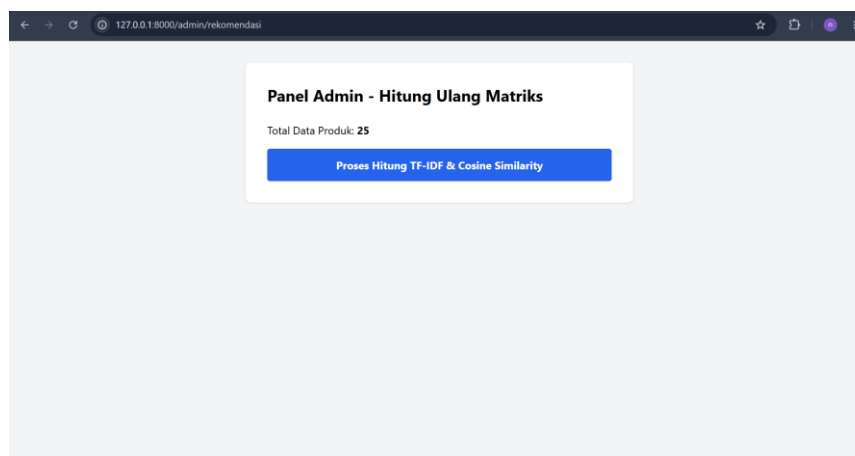
Pada halaman detail, sistem menampilkan informasi lengkap produk serta daftar rekomendasi produk serupa. Produk rekomendasi diurutkan berdasarkan nilai similarity score tertinggi sehingga produk yang paling mirip muncul di urutan pertama.



Gambar 5. Detail dan Rekomendasi

c. Halaman Admin

Halaman admin digunakan untuk menjalankan proses perhitungan ulang TF-IDF dan Cosine Similarity. Admin juga dapat melihat jumlah data produk yang tersedia pada sistem.



Gambar 6. Halaman Admin

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem rekomendasi oleh-oleh kerajinan khas Yogyakarta menggunakan pendekatan Content-Based Filtering (CBF) berhasil dirancang dan diimplementasikan dengan baik menggunakan framework Laravel 9 dan basis data MySQL. Sistem memanfaatkan atribut produk berupa kategori, bahan, lokasi asal, dan fungsi produk sebagai representasi karakteristik setiap produk. Tahapan preprocessing dilakukan untuk menghasilkan data yang lebih terstruktur sebelum dilakukan pembobotan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Selanjutnya, metode Cosine Similarity digunakan untuk menghitung tingkat kemiripan antarproduk sehingga sistem dapat menghasilkan rekomendasi berdasarkan nilai similarity tertinggi. Hasil implementasi menunjukkan bahwa sistem mampu memberikan rekomendasi produk yang relevan tanpa memerlukan data riwayat transaksi maupun penilaian pengguna. Sebagai contoh, produk “Batik Kawung Elegan” berhasil direkomendasikan bersama produk “Batik Tulis Giriloyo” dengan nilai similarity tertinggi sebesar 0,4453, diikuti oleh produk “Syal Batik Motif Sekar Jagad” dengan nilai similarity 0,3316, serta produk “Tas Anyaman Bambu” dan “Celengan Gerabah” dengan nilai similarity masing-masing sebesar 0,0393 dan 0,0377, sehingga menunjukkan sistem mampu mengurutkan rekomendasi berdasarkan tingkat kemiripan atribut produk secara konsisten. Selain itu, penyimpanan hasil perhitungan pada tabel matriks kemiripan membuat proses pencarian rekomendasi menjadi lebih cepat dan efisien karena perhitungan tidak dilakukan berulang setiap kali pengguna mengakses sistem. Antarmuka yang terdiri atas halaman katalog, detail produk, rekomendasi, dan halaman administrasi juga berfungsi dengan baik sehingga mendukung kemudahan penggunaan dan pengelolaan data. Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan Content-Based Filtering yang dipadukan dengan TF-IDF dan Cosine Similarity efektif diterapkan untuk membantu pengguna menemukan oleh-oleh kerajinan khas Yogyakarta yang sesuai dengan karakteristik produk yang dipilih. Sistem ini diharapkan dapat meningkatkan pengalaman pengguna dalam memilih produk sekaligus mendukung promosi dan pemasaran produk UMKM kerajinan lokal secara lebih efektif di era digital.

REFERENCES

- Alamsyah, A. A., Nugraha, R. N., Mulya, M. H., Putra, E. S., & Jhanufa, A. B. I. (2023). Keberagaman budaya Yogyakarta sebagai destinasi wisata budaya unggulan. *Jurnal Ilmiah Wahana Pendidikan*, 9(25), 771–780.
- Anggraeni, Y. S., Saputri, S. D., Azzahra, T., Verrel, A. G., & Tundjungsari, V. (2026). Implementasi Content-Based Filtering Menggunakan TF-IDF dan Cosine Similarity untuk Rekomendasi Buku Akademik Mahasiswa. *Jurnal Informatika: Jurnal Pengembangan IT*, 11(2), 213–225.
- Phalle, T. S., & Bhushan, S. (2024). Content Based Filtering And Collaborative Filtering: A Comparative Study. *Journal of Advanced Zoology*, 45.

- Praditya, N. W. P. Y., Permanasari, A. E., & Hidayah, I. (2021). Designing a tourism recommendation system using a hybrid method (Collaborative Filtering and Content-Based Filtering). *2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, 298–305.
- Pranata, A., & Sulianta, F. (2025). Implementasi dan Analisis Algoritma Content-based filtering pada sistem rekomendasi produk tas pada basis data mysql. *Jurnal Ilmiah Global Education*, 6(3), 1419–1444.
- Rachmaniar, A., Widayati, S., & Rokoyah, K. (2025a). Sistem Rekomendasi Produk E-Commerce menggunakan Collaborative Filtering dan Content-Based Filtering. *Journal of Information System, Informatics and Computing*, 9(1), 40–54.
- Rachmaniar, A., Widayati, S., & Rokoyah, K. (2025b). Sistem Rekomendasi Produk E-Commerce menggunakan Collaborative Filtering dan Content-Based Filtering. *Journal of Information System, Informatics and Computing*, 9(1), 40–54.
- Rahmawati, R., & Suwarni, E. (2023). Ulasan Produk dan Jumlah Produk Terjual Dampaknya pada Keputusan Pembelian di Marketplace Shopee. *Jurnal Digital Bisnis, Modal Manusia, Marketing, Entrepreneurship, Finance, & Strategi Bisnis (Dimensi)*, 3(1), 46–53.
- Rivalni, R., Putri, W. G., Salsyabilla, V. T., & Wardani, W. (2025). Pengembangan Dan Upaya Ekonomi Kreatif Melalui Umkm Dalam Meningkatkan Potensi Pariwisata. *Seminar Nasional Pariwisata Dan Kewirausahaan (SNPK)*, 4, 394–403.
- Siahaan, A. M., Robet, R., Fernando, J. K., & Natalia, D. (2024). Pemanfaatan Dana Desa Untuk Meningkatkan Kesejahteraan Masyarakat Melalui Promosi Produk UMKM Menggunakan Collaborative Filtering Berbasis Android. *Bulletin of Computer Science Research*, 5(1), 24–32.
- Timur, A. Y. (2025). Sistem rekomendasi lagu indonesia menggunakan metode content-based filtering dan cosine similarity. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(1).