



## Sistem Rekomendasi Produk Makeup Berbasis Content-Based Filtering dengan TF-IDF dan Cosine Similarity

Nur Oktavin Idris<sup>1\*</sup>, Fuad Pontoioy<sup>2</sup>

<sup>1</sup>Universitas Negeri Gorontalo, Indonesia, e-mail: nur.oktavin@ung.ac.id

<sup>2</sup>Universitas Negeri Gorontalo, Indonesia, e-mail: fuad.pontoioy@ung.ac.id

### Info Artikel

**Diajukan:** 24-07-2025

**Diterima:** 25-07-xxxx

**Diterbitkan:** 37-07-2025

**Kata Kunci:**

Sistem rekomendasi;

Makeup;

Content-based filtering;

TF-IDF;

Cosine similarity;

**Keywords:**

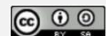
Recommendation system;

Makeup;

Content-based filtering;

TF-IDF;

Cosine similarity;



Lisensi: cc-by-sa

Copyright © 2025 by Author.

Published by Faatuatua Media Karya

### Abstrak

Industri kosmetik yang semakin berkembang menghadirkan beragam produk makeup, namun konsumen sering mengalami kesulitan dalam memilih produk alternatif yang sesuai dengan preferensi mereka. Penelitian ini bertujuan mengembangkan sistem rekomendasi produk makeup berbasis content-based filtering untuk membantu pengguna menemukan produk yang relevan ketika produk utama tidak tersedia. Metode yang digunakan mencakup pengumpulan data dari Kaggle dengan dataset dari Luxxify Makeup, eksplorasi data, prapemrosesan data, serta ekstraksi fitur menggunakan TF-IDF dan perhitungan kemiripan antarproduk dengan cosine similarity. Produk yang direkomendasikan merupakan produk dengan kemiripan tertinggi berdasarkan kategori dan deskripsi produk. Evaluasi dilakukan menggunakan metrik Mean Average Precision (MAP) untuk menilai relevansi rekomendasi. Hasil penelitian menunjukkan bahwa sistem berhasil merekomendasikan lima produk alternatif yang sesuai dengan tingkat akurasi yang sangat tinggi (MAP = 1.00). Sistem ini berkontribusi dalam memberikan solusi pencarian produk secara personal dan efisien, serta dapat diterapkan pada platform e-commerce atau layanan digital kecantikan.

### Abstract

The growing cosmetics industry offers a wide range of makeup products; however, consumers often face difficulties in selecting alternative products that align with their preferences. This study aims to develop a content-based filtering recommendation system to assist users in finding relevant products when their primary product is unavailable. The method includes data collection from Kaggle using the Luxxify Makeup dataset, data exploration, preprocessing, feature extraction using TF-IDF, and product similarity calculation using cosine similarity. The recommended products are those with the highest similarity based on category and product description. Evaluation was carried out using the Mean Average Precision (MAP) metric to assess the relevance of recommendations. The results show that the system successfully recommends five alternative products with very high accuracy (MAP = 1.00). This system contributes to providing a personalized and efficient product search solution and can be applied to e-commerce platforms or digital beauty services.

## 1. PENDAHULUAN

Industri kecantikan, khususnya produk makeup, mengalami pertumbuhan yang signifikan di Indonesia dalam beberapa tahun terakhir. Data menunjukkan bahwa jumlah perusahaan kosmetik meningkat dari 819 pada tahun 2021 menjadi lebih dari 1.000 pada akhir tahun 2023[1]. Pertumbuhan ini didorong oleh maraknya inovasi produk lokal, tren kecantikan yang dinamis, serta peningkatan konsumsi masyarakat terhadap produk perawatan diri[2]. Namun, seiring dengan melimpahnya pilihan produk di pasaran, konsumen kerap mengalami kesulitan dalam memilih produk yang sesuai dengan kebutuhan dan preferensinya. Ketika produk favorit mereka habis atau tidak lagi tersedia, proses pencarian alternatif yang tepat seringkali menjadi tantangan tersendiri, terutama tanpa adanya panduan atau sistem yang membantu secara personal.

Permasalahan utama yang diangkat dalam penelitian ini adalah bagaimana memberikan rekomendasi produk make-up secara personal kepada pengguna berdasarkan preferensi dan kesamaan fitur produk. Beberapa aspek yang menambah kompleksitas masalah ini antara lain adalah keragaman kategori produk, perbedaan brand, variasi harga, serta ulasan konsumen yang sangat subjektif. Dalam konteks ini, pengguna membutuhkan sistem yang tidak hanya dapat menyajikan daftar produk populer, tetapi mampu memahami preferensi mereka dan menyarankan produk yang relevan dari berbagai merek dan kategori secara otomatis.

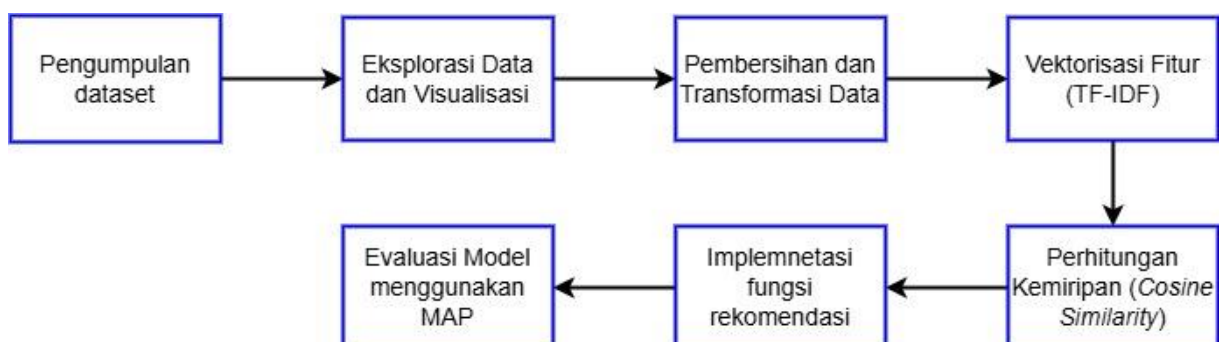
Beberapa penelitian terdahulu telah mengembangkan sistem rekomendasi di bidang kecantikan, seperti menerapkan *content-based filtering* untuk merekomendasikan produk Emina Cosmetics[3]. Namun, sistem ini terbatas pada satu merek dan belum mengakomodasi keberagaman produk lintas brand. Penelitian lainnya juga mengembangkan sistem serupa untuk pemilihan produk skincare[4], tetapi tidak menyertakan evaluasi kuantitatif berbasis metrik performa rekomendasi seperti *Mean Average Precision* (MAP). Perbedaan mendasar dari penelitian ini dibandingkan dengan studi sebelumnya adalah penerapan pendekatan *content-based filtering* yang diperkuat dengan representasi fitur produk menggunakan teknik TF-IDF (*Term Frequency-Inverse Document Frequency*) serta pengukuran kemiripan antarproduk melalui algoritma *cosine similarity*[5]. Inovasi ini memungkinkan sistem untuk secara lebih akurat mengenali kesamaan antarproduk berdasarkan kategori dan deskripsi, sehingga mampu merekomendasikan alternatif produk lintas merek yang lebih relevan dan personal.

Tujuan dari penelitian ini adalah membangun sistem rekomendasi produk make-up yang mampu menyarankan top-N produk berdasarkan fitur produk seperti kategori, brand, dan deskripsi. Sistem dikembangkan menggunakan teknik *content-based filtering* yang diimplementasikan melalui representasi teks menggunakan TF-IDF dan perhitungan kemiripan dengan *cosine similarity*. Evaluasi performa sistem dilakukan menggunakan metrik MAP untuk memastikan bahwa hasil rekomendasi tidak hanya relevan tetapi juga tersusun dalam urutan yang optimal[6].

Inovasi utama dalam penelitian ini terletak pada integrasi preprocessing berbasis vektorisasi fitur tekstual produk dan pemodelan kemiripan antarproduk yang tidak bergantung pada perilaku historis pengguna, sehingga tetap efektif meskipun dalam kondisi *cold start*. Dengan memanfaatkan dataset public (Kaggle) dari Luxxify Makeup yang terdiri dari lebih dari seribu produk dari berbagai merek, penelitian ini memberikan kontribusi baru terhadap pengembangan sistem rekomendasi dalam industri kecantikan yang lebih inklusif, adaptif, dan mampu menangani kebutuhan pengguna dalam konteks yang lebih luas dibandingkan pendekatan-pendekatan sebelumnya.

## 2. METODE PENELITIAN

Penelitian ini merupakan penelitian terapan yang menggunakan pendekatan kuantitatif dalam pengembangan sistem Rekomendasi Berbasis konten (*content-based filtering*). Penelitian ini dirancang secara eksperimental dengan tujuan membangun dan menguji sistem yang mampu merekomendasikan produk make-up berdasarkan kesamaan fitur deskriptif produk. Desain penelitian mengadopsi alur pengembangan sistem machine learning yang mencakup tahap eksplorasi data, preprocessing, pemodelan, dan evaluasi performa sistem. yang dipergunakan selama proses penelitian. Berikut alur atau langkah-langkah penelitian yang ditunjukkan pada gambar 1.



**Gambar 1.** Alur penelitian

### 2.1 Pengumpulan dan Pemahaman Data

Dataset yang digunakan adalah *Luxxify Makeup Dataset* yang diunduh dari platform Kaggle. Dataset terdiri atas dua file utama:

1. *cleaned\_makeup\_products.csv* : berisi deskripsi produk makeup (jumlah 1373 entri, 37 atribut )
2. *cleaned\_makeup\_reviews.csv* : untuk dataset ini tidak digunakan dalam model rekomendasi untuk penelitian ini.

Penelitian ini difokuskan pada data produk, khususnya atribut seperti: *product\_name*, *category*, *brand*, *price*, *description*, *pros*, *cons*, *best\_uses*, dan *average\_rating*.

## 2.2 Eksplorasi Data Awal (EDA)

Penelitian dilakukan dengan eksplorasi data melalui analisis statistic dan visualisasi berikut:

1. Distribusi produk per kategori dan merek
2. Distribusi rating produk dan ulasan pengguna
3. Identifikasi nilai hilang (missing values) dan duplikasi data.

Dalam menghitung jumlah missing values di setiap kolom dalam dataset menggunakan fungsi `products_relevant.isnull().sum()`. Analisis ini membantu mengidentifikasi kolom mana yang memerlukan penanganan, seperti imputasi atau penghapusan baris/kolom. Langkah ini penting untuk memastikan dataset siap digunakan untuk analisis dan pemodelan.

## 2.3 Prapemrosesan Data

Tahap ini bertujuan membersihkan dan menyiapkan data sebelum pemodelan:

1. Pemilihan atribut yang relevan untuk sistem rekomendasi

Dataset biasanya memiliki banyak kolom, tetapi tidak semuanya relevan untuk analisis atau pemodelan. Untuk itu pada model ini, hanya memilih kolom yang relevan dari file detail produk untuk membantu mengurangi kompleksitas data dan meningkatkan efisiensi analisis lebih lanjut pada sistem rekomendasi. Kolom yang dipilih mencakup *product\_link\_id* sebagai ID unik produk, *product\_name* untuk nama produk, *category* untuk kategori produk, dan *brand* untuk merek. Selain itu, kolom *price* menunjukkan harga produk, sementara *description* memberikan deskripsi produk. Kolom seperti *pros*, *cons*, dan *best\_uses* digunakan untuk menganalisis umpan balik pengguna tentang kelebihan, kekurangan, serta penggunaan terbaik produk. Terakhir, kolom *average\_rating* memberikan gambaran umum tentang kualitas produk berdasarkan rata-rata rating dari pengguna.

2. Penghapusan baris yang mengandung *missing values*

Untuk mengatasi missing value digunakan fungsi `dropna()` untuk membersihkan missing value dari dataset subset `products_relevant`. Fungsi ini akan menghapus semua baris yang memiliki nilai kosong (missing values) di salah satu kolom yang dipilih. Hasilnya adalah DataFrame baru bernama `product_clean`, yang hanya berisi data lengkap tanpa nilai kosong. Meskipun metode ini efektif untuk memastikan integritas data, penggunaan `dropna()` dapat menyebabkan hilangnya banyak data jika terdapat banyak missing value dalam dataset, terutama pada kolom yang jarang terisi. Oleh karena itu, metode ini cocok digunakan jika data yang hilang relatif sedikit atau jika data yang lengkap lebih penting untuk analisis dibandingkan mempertahankan jumlah data yang besar.

3. Penghapusan duplikasi data

4. Perubahan struktur data ke dalam format yang dapat diproses oleh algoritma TF-IDF

## 2.4 Ekstraksi Fitur Teks dengan TF-IDF

Teknik *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk mengubah data kategori produk ke dalam bentuk vektor numerik. Proses ini menciptakan matriks fitur untuk mewakili setiap produk berdasarkan deskripsi kategorinya, sehingga sistem dapat menghitung kesamaan antarproduk. Nilai TF-IDF diperoleh melalui perhitungan Term Frequency (TF), yang merepresentasikan seberapa sering suatu istilah subjek muncul dalam sebuah dokumen, serta Inverse Document Frequency (IDF), yang menunjukkan seberapa banyak dokumen dalam koleksi yang memuat istilah subjek tersebut. Nilai TF untuk suatu istilah dalam dokumen dapat dihitung menggunakan rumus berikut[7]:

$$tf(t, d) = 1 + \log(f_{td}) \quad (1)$$

Dimana:

$$\begin{aligned} tf(t, d) &= \text{Frekuensi kemunculan istilah } t \text{ dalam dokumen } d \\ f_{td} &= \text{Banyaknya istilah } t \text{ yang terdapat dalam dokumen } d \end{aligned}$$

Perhitungan nilai IDF untuk suatu istilah dalam kumpulan dokumen dapat diperoleh melalui perhitungan menggunakan rumus berikut:

$$idf(t) = \log\left(\frac{N}{df_t}\right) \quad (2)$$

Dimana:

$idf(t)$  = Besaran IDF untuk istilah t dalam kumpulan dokumen  
 $df_t$  = Jumlah dokumen yang mengandung istilah t  
 $N$  = Total dokumen yang memuat istilah t

Perhitungan bobot TF-IDF untuk istilah t pada dokumen d dilakukan dengan rumus berikut:

$$tf - idf(t, d) = tf(t, d) * idf(t) \quad (3)$$

Dimana:

$tf - idf(t, d)$  = Bobot TF-IDF dari istilah t terhadap dokumen d dalam koleksi dokumen  
 $tf(t, d)$  = Nilai TF dari istilah t pada dokumen d  
 $idf(t)$  = Nilai IDF dari istilah t dalam kumpulan dokumen

## 2.5 Perhitungan Similaritas dengan Cosine Similarity

Cosine similarity adalah metode untuk mengukur seberapa mirip dua vektor dalam ruang multidimensi. Ini adalah pengukuran kosinus sudut antara dua vektor yang dimensi dan magnitudonya direpresentasikan sebagai titik dalam ruang. Nilai similaritas kosinus berkisar antara -1 hingga 1, di mana nilai 1 menunjukkan kedua vektor sepenuhnya sejajar (100% mirip), 0 menunjukkan vektor tegak lurus (tidak ada keterkaitan), dan -1 menunjukkan kedua vektor sepenuhnya berlawanan arah (100% tidak mirip). Metode ini sering digunakan dalam pemrosesan teks dan pengelompokan data untuk menentukan tingkat kesamaan antara dokumen atau fitur dalam dataset[5]. *Cosine similarity* dituliskan dalam rumus[8]:

$$CosineSimilarity(A_k, B_k) = \frac{\sum_k(A_k * B_k)}{\sqrt{\sum_k A_k^2} * \sqrt{\sum_k B_k^2}} \quad (4)$$

Setelah mendapatkan representasi numerik dari setiap produk, dilakukan perhitungan derajat kemiripan antarproduk menggunakan *cosine similarity*. Nilai *similarity* ini menunjukkan sejauh mana dua produk memiliki kemiripan berdasarkan kontennya.

## 2.6 Pengembangan Fungsi Rekomendasi

Dibuat fungsi rekomendasi\_produk() untuk menampilkan Top-N produk yang paling mirip dengan produk pilihan pengguna. Fungsi ini akan:

1. Mengurutkan produk berdasarkan skor kemiripan
2. Menghapus produk input agar tidak direkomendasikan kembali
3. Menggabungkan informasi nama produk, kategori, dan merek

## 2.7 Evaluasi Model

Evaluasi sistem dilakukan dengan menggunakan metrik Mean Average Precision (MAP) untuk menilai relevansi dan urutan hasil rekomendasi.

Langkah-langkah pengujian:

1. Pengujian Produk Acuan: Dipilih produk "Bio Stick Foundation" sebagai item input.
2. Penerapan Model: Fungsi rekomendasi\_produk() dijalankan untuk menghasilkan 5 produk teratas berdasarkan kemiripan kategori.
3. Validasi Hasil: Daftar rekomendasi dibandingkan dengan semua produk dalam kategori *Foundation* sebagai daftar item relevan.
4. Perhitungan MAP: Dihitung seberapa akurat sistem menempatkan produk yang relevan di posisi atas rekomendasi.

Mean Average Precision (MAP) pada proyek ini digunakan untuk mengukur seberapa baik sistem rekomendasi dalam menampilkan item yang relevan bagi pengguna. MAP menilai tidak hanya apakah item yang direkomendasikan relevan, tetapi juga posisi (ranking) item tersebut dalam daftar rekomendasi. MAP membantu mengevaluasi apakah sistem ini benar-benar merekomendasikan item yang paling relevan berdasarkan preferensi sebelumnya. MAP dihitung berdasarkan Average Precision (AP), yaitu rata-rata precision untuk setiap item relevan dalam daftar rekomendasi, dengan nilai 0 diberikan jika item relevan tidak muncul. Perhitungan precision dalam AP mempertimbangkan urutan

item yang diberikan oleh sistem, sehingga setiap item relevan yang muncul dalam daftar rekomendasi memiliki nilai precision tersendiri. Rumus MAP berikut[6]:

$$MAP(Q) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \frac{1}{m_j} \sum_{k=1}^{m_j} Precision(R_{jk})$$

Dimana:

Q : jumlah *query* uji

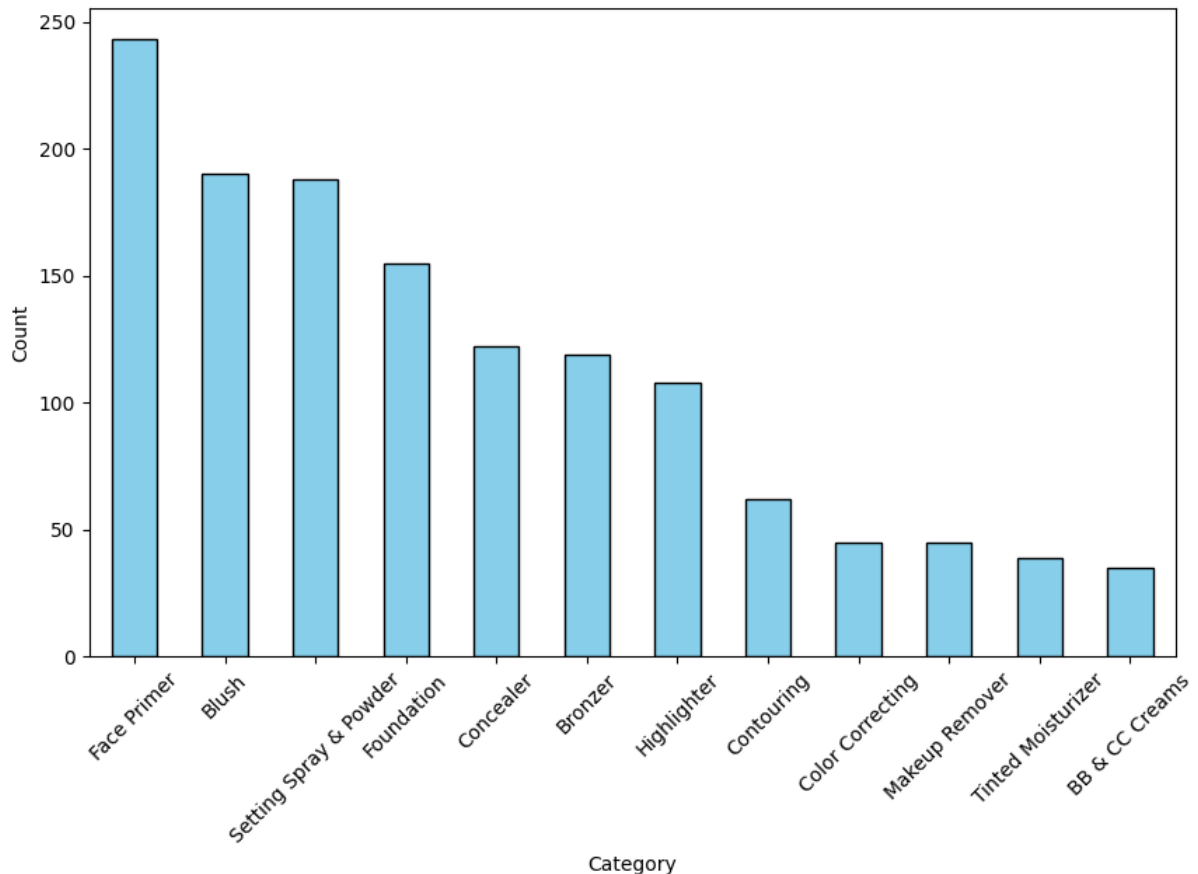
R : item relevan yang dihasilkan oleh sistem

m : jumlah item relevan yang dihasilkan dari *query*

Perhitungan precision dalam AP mempertimbangkan urutan item yang diberikan oleh sistem, sehingga setiap item relevan yang muncul dalam daftar rekomendasi memiliki nilai precision tersendiri. Semakin tinggi MAP, semakin baik sistem dalam memberikan rekomendasi yang relevan dan terurut dengan optimal.

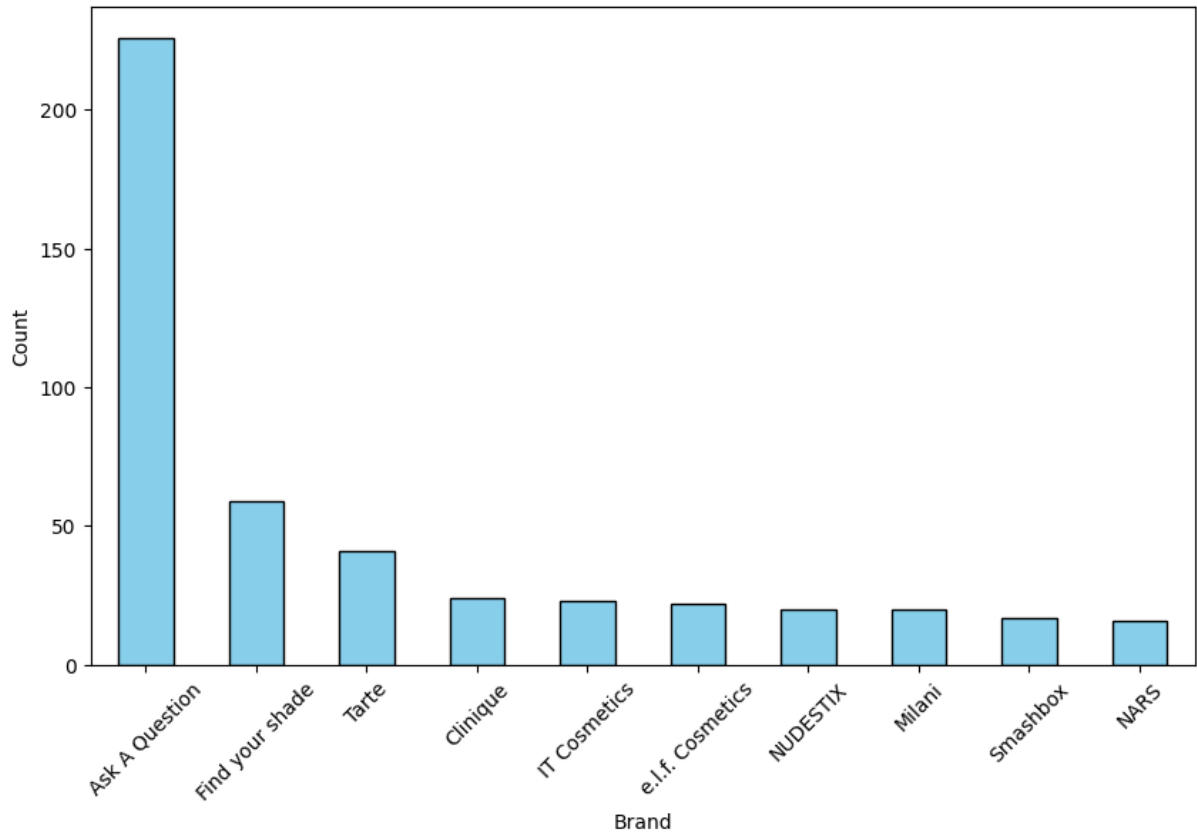
### 3. HASIL DAN ANALISIS

#### 3.1 Eksplorasi Data Awal (EDA)



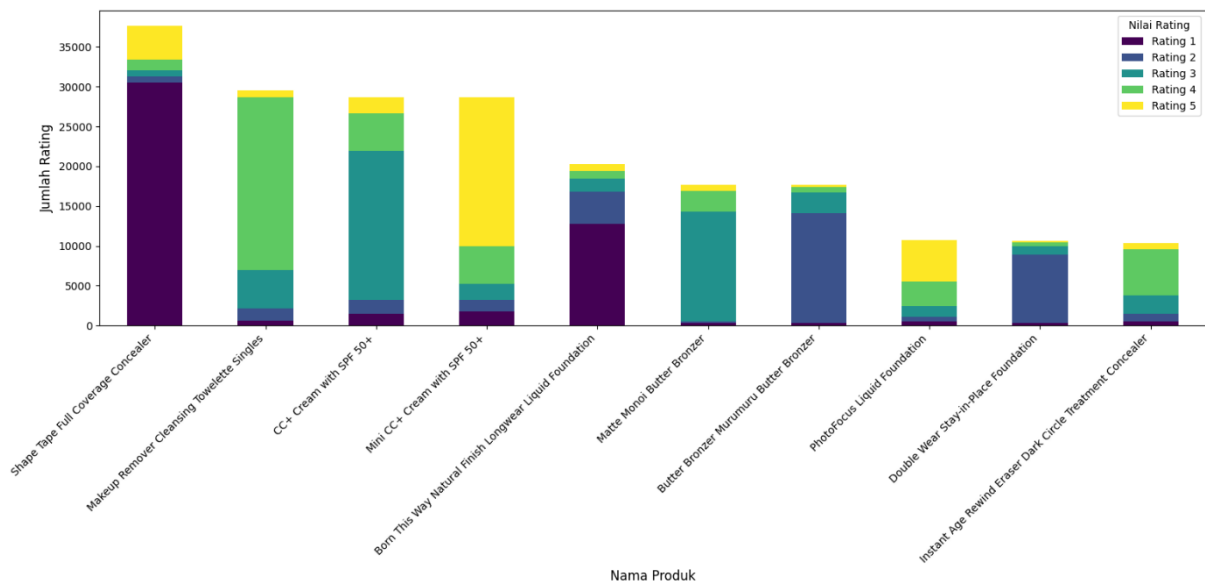
**Gambar 2.** Distribusi produk tiap kategori

Berdasarkan gambar 2, dapat disimpulkan bahwa distribusi produk makeup menunjukkan bahwa kategori *Face Primer* memiliki jumlah produk terbanyak dibandingkan dengan kategori lainnya, diikuti oleh *Blush*, *Setting Spray & Powder*, dan *Foundation*. Sebaliknya, kategori dengan jumlah produk paling sedikit adalah *BB & CC Creams*. Hal ini mengindikasikan bahwa *Face Primer* merupakan salah satu kategori yang paling banyak dikembangkan atau dipasarkan oleh produsen dalam dataset ini.



**Gambar 3.** Distribusi produk tiap merek

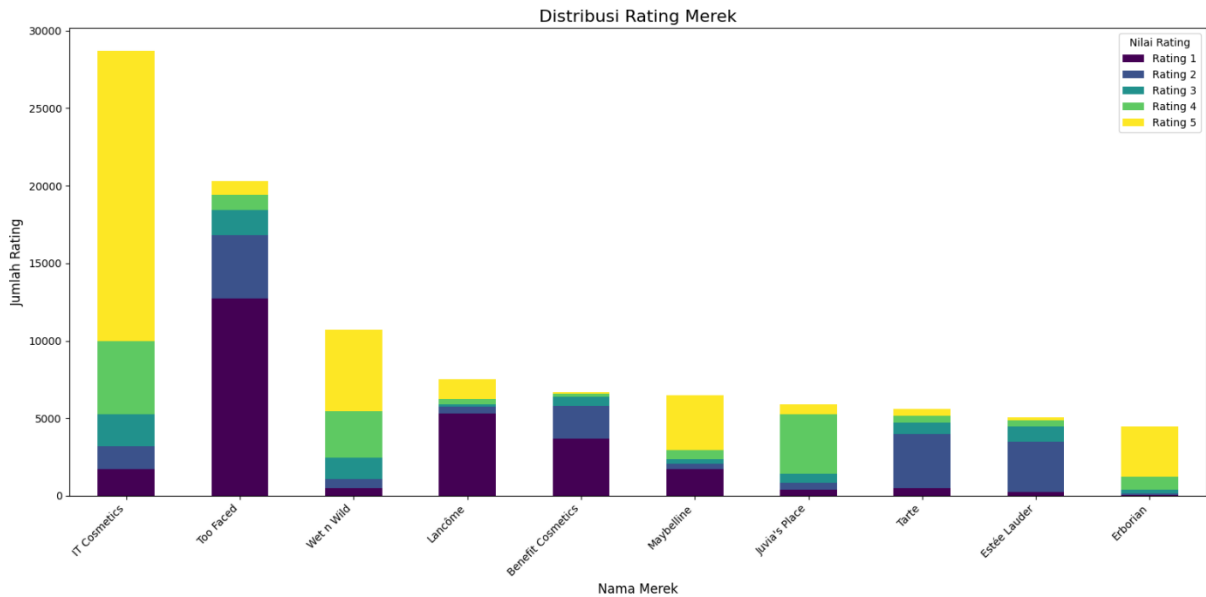
Berdasarkan grafik distribusi produk tiap merek, terlihat bahwa merek *Ask A Question* memiliki jumlah produk yang jauh lebih banyak dibandingkan dengan merek lainnya, sedangkan merek-merek seperti *NARS*, *Smashbox*, dan *Milani* memiliki jumlah produk yang jauh lebih sedikit. Hal ini menunjukkan dominasi satu merek tertentu di dataset, sementara distribusi produk untuk merek lainnya relatif merata dengan jumlah yang lebih kecil.



**Gambar 4.** Distribusi rating produk

Berdasarkan grafik distribusi rating produk, terlihat bahwa produk *Shape Tape Full Coverage Concealer* memiliki jumlah total rating tertinggi dibandingkan produk lainnya. Sebagian besar rating untuk produk-produk yang ditampilkan cenderung didominasi oleh rating tinggi, terutama Rating 4 dan

Rating 5, yang menunjukkan kepuasan pengguna secara umum terhadap produk-produk tersebut. Namun, ada juga kontribusi kecil dari Rating 1 dan Rating 2 di semua produk, yang menunjukkan beberapa ketidakpuasan pelanggan. Produk-produk dengan rating tertinggi menunjukkan potensi penerimaan pasar yang baik, sementara yang lain memiliki peluang untuk peningkatan kualitas.



Gambar 5. Distribusi rating merek

Berdasarkan grafik distribusi rating merek, terlihat bahwa merek *IT Cosmetics* memiliki jumlah rating tertinggi dibandingkan merek-merek lainnya, dengan dominasi Rating 5 dan Rating 4 yang menunjukkan tingkat kepuasan pelanggan yang tinggi terhadap produk mereka. Merek *Too Faced* juga memiliki jumlah rating yang signifikan, namun terlihat lebih beragam dengan kontribusi yang cukup besar dari rating yang lebih rendah seperti Rating 1 dan Rating 2. Merek-merek lainnya, seperti *Wet n Wild*, *Lancôme*, dan *Benefit Cosmetics*, memiliki jumlah rating yang lebih sedikit namun tetap didominasi oleh Rating 4 dan Rating 5, menunjukkan kecenderungan positif dalam kepuasan pelanggan secara keseluruhan. Hal ini mencerminkan bahwa sebagian besar merek di grafik ini memiliki persepsi yang baik di pasar, terutama pada produk-produk unggulan mereka.

### 3.2 Hasil rekomendasi

Dalam mendapatkan rekomendasi diilustrasikan seperti berikut, dimana pengguna X pernah membeli produk "Bio Stick Foundation". Kemudian, saat pengguna tersebut berencana untuk membeli produk dari merk lain, sistem akan merekomendasikan produk yang mirip hasil rekomendasi produk yang sebelumnya digunakan oleh pengguna X.

Rekomendasi produk ini berdasarkan kesamaan yang dihitung dengan cosin similarity pada tahap sebelumnya. Parameter yang digunakan untuk membuat fungsi rekomendasi produk yaitu: Nama\_produk, similarity\_data, items (produk, kategori), dan k (banyak rekomendasi yang diberikan). Keluaran sistem ini adalah berupa top-N recommendation. Sehingga akan memberikan sejumlah rekomendasi produk pada pengguna yang diatur dalam parameter k. Dalam kasus ini, nantinya akan mencari produk yang mirip dengan "Bio Stick Foundation", sehingga perlu drop nama\_produk "Bio Stick Foundation" agar tidak muncul dalam daftar rekomendasi yang diberikan dan menemukan rekomendasi produk yang mirip dengan "Bio Stick Foundation" sehingga akan muncul tampilan seperti tabel 1.

**Tabel 1.** Hasil rekomendasi produk

	produk	kategori	merek
0	TRUE COLOR Skin Perfecting Stick Foundation SP...	Foundation	BLK/OPL
1	Flawless Creator Liquid Foundation Drops	Foundation	Dermablend
2	PHYTO-PIGMENTS Flawless Serum Foundation	Foundation	Juice Beauty
3	Light Reflecting Advanced Skincare Foundation	Foundation	NARS
4	Conceal & Define Full Coverage Foundation	Foundation	Revolution Beauty

Berdasarkan tabel 1, untuk hasil pengujian model Content-Based Filtering, sistem berhasil menyajikan rekomendasi top 5 produk yang mirip dengan “Bio Stick Foundation”, semuanya termasuk dalam kategori foundation dari berbagai merek. Hal ini menunjukkan bahwa sistem mampu mengidentifikasi produk serupa berdasarkan kemiripan fitur kategori. Pendekatan ini memungkinkan pengguna yang menyukai “Bio Stick Foundation” untuk menemukan alternatif produk lain yang sesuai dengan preferensi mereka dalam kategori yang sama, memberikan pengalaman rekomendasi yang relevan dan bervariasi.

### 3.3 Evaluasi Model

Metrik evaluasi digunakan untuk menilai seberapa baik performa suatu model. Dalam konteks ini, metrik evaluasi yang digunakan untuk mengukur kinerja model yaitu Mean Average Precision (MAP). Metrik ini bertujuan untuk memberikan gambaran tentang seberapa baik model bekerja dalam merekomendasikan item berdasarkan preferensi pengguna.

Pada penelitian ini diperoleh hasil perhitungan Mean Average Precision (MAP) = 1.00 yang menunjukkan bahwa sistem rekomendasi memiliki performa yang sangat baik dalam memberikan rekomendasi item yang relevan kepada pengguna, dan menandakan bahwa semua item yang direkomendasikan oleh model benar-benar relevan, dan urutan rekomendasinya juga optimal, di mana item relevan muncul di posisi awal daftar rekomendasi. Namun, dalam praktiknya, MAP yang terlalu tinggi bisa menjadi indikasi overfitting atau kurangnya variasi dalam dataset, sehingga perlu dilakukan evaluasi lebih lanjut dengan dataset yang lebih beragam untuk memastikan bahwa sistem dapat bekerja dengan baik di berbagai skenario pengguna.

## 4. KESIMPULAN

Sistem rekomendasi produk makeup yang dikembangkan menggunakan Content-Based Filtering berhasil memberikan rekomendasi yang relevan dan bermanfaat bagi pengguna. Pendekatan ini mampu menyajikan rekomendasi spesifik berdasarkan preferensi pengguna. Evaluasi model memberikan dampak signifikan dalam penelitian sistem rekomendasi ini. Hal ini mendukung efisiensi sistem dalam merekomendasikan produk yang relevan, yang secara langsung berdampak pada pengalaman pengguna. Dalam konteks bisnis, evaluasi ini menegaskan bahwa sistem dapat secara andal mengelompokkan produk berdasarkan fitur yang serupa, seperti kategori. Hasilnya adalah rekomendasi yang lebih relevan dan personal, yang dapat meningkatkan kepuasan pengguna, loyalitas pelanggan, dan konversi penjualan. Dengan kemampuan model yang terbukti efektif, perusahaan brand produk dapat membangun kepercayaan pada sistem rekomendasi untuk mendukung pertumbuhan bisnis, terutama di pasar produk kecantikan Indonesia yang kompetitif. Evaluasi yang baik juga memberikan kepercayaan kepada pemangku kepentingan untuk mengintegrasikan sistem ini ke dalam strategi bisnis yang lebih luas.

## REFERENSI

- [1] M. F. Milenao, “Industri Kosmetik Lokal Kian Meroket, Pertumbuhan Tembus Angka 48 Persen,” [goodnewsfromindonesia.id](https://www.goodnewsfromindonesia.id). Accessed: Jul. 22, 2025. [Online]. Available: <https://www.goodnewsfromindonesia.id/2024/08/22/industri-kosmetik-lokal-kian-meroket-pertumbuhan-tembus-angka-48-persen>
- [2] Warta Ekonomi, “Tahan Banting! Industri Kecantikan Indonesia Makin Potensial,” [wartaekonomi.co.id](https://wartaekonomi.co.id). Accessed: Jul. 22, 2025. [Online]. Available: <https://wartaekonomi.co.id/read554079/tahan-banting-industri->

- kecantikan-indonesia-makin-potensial
- [3] F. B. A. Larasati and H. Februiyanti, "Sistem Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode Content - Based Filtering," *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 45–54, 2021, doi: 10.36595/misi.v4i1.250.
  - [4] A. Sulami, V. Atina, and N. Numalitasari, "Penerapan Metode Content Based Filtering dalam Sistem Rekomendasi Pemilihan Produk Skincare," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 9, no. 2, pp. 172–181, 2024, doi: 10.30998/string.v9i2.24066.
  - [5] S. Seelam, "Machine Learning Fundamentals: Cosine Similarity and Cosine Distance," medium.com. Accessed: Jul. 22, 2025. [Online]. Available: <https://medium.com/geekculture/cosine-similarity-and-cosine-distance-48eed889a5c4>
  - [6] W. G. Suka Parwita and E. Winarko, "Hybrid Recommendation System Memanfaatkan Penggalian Frequent Itemset dan Perbandingan Keyword," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 9, no. 2, pp. 167–176, 2015, doi: 10.22146/ijccs.7545.
  - [7] H. H. Arfisko and A. T. Wibowo, "Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering dan Content-Based Filtering," *e-Proceeding Eng.*, vol. 9, no. 3, pp. 2149–2159, 2022.
  - [8] Supiyanto and Sriyono, "Metode Cosine Similarity Untuk Mendeteksi Kemiripan Pada Dokumen Teks," *SAINS J. MIPA dan Pengajarannya*, vol. 1, no. 1, pp. 1–7, 2023, [Online]. Available: <https://ejournal.uncen.ac.id/index.php/SAINS>