

Penggunaan Metode Content Based Filtering Dalam Sistem Rekomendasi Elektronik

Ester Anugrayningtyas¹, Joni Mulindar², Herliyani Hasanah³

¹²³Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Duta Bangsa Surakarta

¹202030324@mhs.udb.ac.id *, ²joni_maulindar@udb.ac.id, ³herliyani_hasanah@udb.ac.id

Abstract

The study aims to develop a more personal electronic product recommendation system according to consumer preference. The methods used are advanced -based filtering through a frequency of frequency of frequency frequency documentations (tf-idf) and cosine documentations. The needed data consists of product information such as name, brand, type, price, and discount. Analysis shows, the ice 1 door of the pre-18 polytron brands at rp 2,100,000 has the highest level of similarity (64.86%) to the 1 door of the polytron type PRW 25. Similar products are refrigerators 1 door polytron type PRB 159 (62,53%) and PRB type 177 (52,55%). Although the 2 doors of the polytron type PRW 25 (\$8.00,000) brands of polytron have the lowest level of similarity (48.99 percent), the product remains relevant as an alternative recommendation. A recommendation system can display alternative products when consumer items are not available. The use of tf-idf and cosine methods has proven to improve the accuracy and personalization of recommendation systems in electronics stores.

Keywords: recommendation system, content based filtering, electronic, TF-IDF, similarity

Abstrak

Penelitian ini bertujuan mengembangkan sistem rekomendasi produk elektronik yang lebih personal sesuai preferensi konsumen. Metode yang digunakan adalah Content-Based Filtering dengan Term Frequency-Inverse Document Frequency (TF-IDF) dan Cosine Similarity. Data yang dibutuhkan berupa informasi produk seperti nama, merek, tipe, harga, dan diskon. Hasil analisis menunjukkan, lemari es 1 pintu merek Polytron tipe PRA 18 dengan harga Rp 2.100.000 memiliki tingkat kesamaan tertinggi (64,86%) dengan lemari es 1 pintu merek Polytron tipe PRW 25. Produk lain yang mirip adalah lemari es 1 pintu merek Polytron tipe PRB 159 (62,53%) dan tipe PRB 177 (52,55%). Meskipun lemari es 2 pintu merek Polytron tipe PRW 25 (Rp 4.700.000) memiliki tingkat kesamaan terendah (48,99%), produk ini tetap relevan sebagai alternatif rekomendasi. Sistem rekomendasi dapat menampilkan produk alternatif ketika barang yang dicari konsumen tidak tersedia. Penggunaan metode TF-IDF dan Cosine Similarity terbukti meningkatkan akurasi dan personalisasi sistem rekomendasi di toko elektronik.

Kata kunci: sistem rekomendasi, content based filtering, elektronik, TF-IDF, similaritas

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International

Pendahuluan

Toko elektronik yang menjual berbagai jenis barang elektronik dengan beragam spesifikasi dapat memberi banyak penawaran barang elektronik pada konsumen toko. Disaat ada konsumen yang ingin membeli barang elektronik akan menanyakan pada karyawan toko mengenai barang dengan spesifikasi yang diinginkan ada. Karyawan pun harus dapat memberi informasi pada konsumen mengenai ketersediaan barang yang diinginkan konsumen. Disaat itu juga karyawan harus dapat memberi rekomendasi beberapa produk yang mirip dengan spesifikasi barang yang diinginkan konsumen.

Banyak kejadian juga dimana konsumen akan langsung menanyakan barang yang mereka inginkan namun dengan batas biaya yang konsumen miliki. Disaat tersebut juga karyawan dapat memberi rekomendasi barang sesuai dengan dana yang konsumen miliki serta mirip dengan spesifikasi barang yang konsumen inginkan.

Adanya persaingan bisnis penjualan elektronik juga dapat terpengaruh dengan kualitas pelayanan dari toko elektronik tersebut pada pelanggan yang harus memberi pengalaman transaksi membeli barang elektronik dengan baik pada konsumen. Untuk memperbaiki kualitas pelayanan karyawan pada konsumen maka peneliti membuat sistem rekomendasi berbasis konten yang dapat digunakan untuk merekomendasikan produk pada konsumen.

Sebuah sistem rekomendasi telah dikembangkan untuk memprediksi dan menyarankan item atau konten yang mungkin diminati oleh pengguna. Sistem ini menggunakan algoritma berbasis data untuk menganalisis pola perilaku pengguna dan preferensi mereka terhadap item yang relevan [1]. Metode Content Based Filtering yang diterapkan pada sistem ini dapat menganalisis deskripsi produk dan menghitung nilai kemiripannya sehingga produk dapat diurutkan berdasarkan nilai kemiripannya [7].

Melalui penelitian yang telah dilakukan oleh Donny Susianto dan Ziad Rusdi [5] yang membuat Sistem

Rekomendasi Pada Penjualan Elektronik Menggunakan Metode *Collaborative Filtering* yang bertujuan untuk membuat sistem agar memberi rekomendasi produk elektronik berdasarkan preferensi pengguna lain karena menganggap bahwa selera pengguna terhadap suatu item akan cenderung sama dari waktu ke waktu, pengguna yang menyukai suatu item biasanya juga akan menyukai item lain dari pengguna lain yang mempunyai kesamaan dengan pengguna tersebut, perbedaan minat konsumen dapat dijadikan informasi bagi konsumen lainnya. Hasil penelitian berupa sistem penjualan elektronik menggunakan metode *collaborative filtering* dapat membantu dalam pemberian rekomendasi produk untuk pengguna. Pengujian sistem ini dilakukan mendapatkan presentase sebesar 64% yang berarti interpretasi skor kategori baik.

Penggunaan metode *collaborative filtering* memberikan rekomendasi produk terbatas pada preferensi pengguna lain dan pengguna itu sendiri, maka penulis melakukan pengembangan sistem dengan menggunakan metode *Content Based Filtering*, dengan melibatkan analisis fitur produk elektronik, metode ini memberikan rekomendasi yang lebih relevan dan sesuai dengan kebutuhan konsumen.

Banyaknya konsumen toko elektronik yang mencari produk elektronik berdasarkan spesifikasi barang yang dibutuhkan serta kesesuaian dengan *budget* yang mereka miliki. Ketika harus menunjukan barang yang diinginkan konsumen maka karyawan toko dapat memberi rekomendasi barang yang mirip dengan yang diinginkan konsumen karena konsumen dapat mempertimbangkan barang yang dapat mereka beli berdasarkan rekomendasi produk yang diberikan konsumen.

Saat karyawan hendak memberi rekomendasi barang elektronik mereka harus dapat menghafal semua produk yang ada pada toko elektronik, dan hal tersebut sering menjadi masalah bagi karyawan toko karena keterbatasan dalam mengingat data produk yang cukup banyak sedangkan konsumen perlu mendapat rekomendasi produk sesegera dalam waktu singkat.

Ketika akan memberikan rekomendasi produk, karyawan juga perlu memperhatikan pembaharuan harga dan ketersediaan barang pada toko elektronik. Banyak kejadian dimana saat ingin memberi rekomendasi produk karyawan harus bertanya pada admin toko mengenai *update* harga barang dan ketersediaan barang terutama ketika produk yang diinginkan konsumen tidak ada maka karyawan harus memberi rekomendasi produk elektronik lain yang mirip dengan yang diinginkan konsumen. Saat harus menghubungi admin tersebut cukup memakan waktu karena karyawan harus menunggu konfirmasi dari pihak admin mengenai barang yang dicari.

Dengan dibuatnya sistem rekomendasi berbasis konten ini mengurangi ketergantungan pada preferensi pengguna lain dan fokus pada karakteristik produk sehingga memberikan rekomendasi yang lebih akurat kepada pengguna berdasarkan spesifikasi barang dan biaya yang mereka inginkan. Sistem rekomendasi berbasis konten dapat membantu karyawan dalam pengecekan data produk yang dicari konsumen dan dapat memberi rekomendasi produk yang kemiripannya tinggi dengan produk yang diinginkan konsumen tanpa memakan banyak waktu [12]. Penelitian ini diharapkan dapat memperbaiki kualitas layanan toko elektronik dalam memberi rekomendasi produk dan memperluas penggunaan metode *Content-Based Filtering* pada konteks penjualan produk elektronik.

Metode Penelitian

Penelitian yang dilakukan oleh peneliti melalui beberapa tahapan berikut:



Gambar 1. Alur Penelitian

1. Identifikasi Masalah dan Tujuan

Peneliti menemukan permasalahan pada sistem pelayanan karyawan pada konsumen yang memakan waktu untuk memberi rekomendasi barang pada konsumen terutama disaat barang yang diinginkan tidak sesuai dengan yang diharapkan oleh konsumen dan ketika barang yang diinginkan konsumen tidak tersedia maka karyawan harus sesegera mungkin memberikan rekomendasi produk yang mirip dengan preferensi konsumen. Maka dengan sistem rekomendasi berbasis konten ini karyawan dapat memberi rekomendasi produk elektronik pada konsumen dengan mudah dan tidak menghambat waktu.

2. Peninjauan Literatur

Untuk membangun sistem rekomendasi, peneliti perlu melakukan riset sistem rekomendasi elektronik yang didapat dari penelitian terdahulu yang dilakukan oleh

Donny Susianto dan Zyad Rusdi [5]. Pada jurnal tersebut peneliti membangun sistem rekomendasi elektronik menggunakan metode *collaborative filtering* yang memberikan rekomendasi berdasarkan preferensi pengguna lain dan pengguna itu sendiri berdasarkan rating yang diberikan pada produk. Peneliti melakukan pengembangan dari sistem tersebut dengan membuat sistem rekomendasi berbasis konten untuk memberi rekomendasi berdasarkan preferensi konsumen sehingga konsumen mendapat rekomendasi barang sesuai dengan yang mereka inginkan dan butuhkan.

3. Pengumpulan Data

Pengumpulan data dilakukan melalui metode observasi langsung di toko salah satu toko elektronik, dengan mengamati dan menganalisis proses serta masalah yang ada. Selain itu, peneliti juga mengumpulkan data historis melalui dokumentasi yang mengacu pada informasi atau data yang telah ada sebelumnya. Data yang dikumpulkan berisi daftar barang elektronik yang diperoleh secara internal dari toko.

4. Pra-Pemrosesan Data

Peneliti menganalisis data produk elektronik yang ada di salah satu toko elektronik, termasuk nama barang, merek, tipe produk, harga, dan diskon. Tujuan dari analisis ini adalah untuk menemukan informasi yang dapat digunakan dalam pembuatan dataset untuk sistem rekomendasi penjualan barang elektronik. Peneliti menggunakan data primer yang dikumpulkan sendiri [2], yaitu informasi mengenai barang elektronik yang dijual untuk direkomendasikan kepada konsumen. Menghapus data-data yang tidak memenuhi atribut atau kriteria yang telah ditentukan [11]. Pembersihan data ini dilakukan guna memberi kemudahan bagi sistem untuk mengidentifikasi setiap item yang ada pada data yang digunakan.

5. Pengembangan Metode Content Based Filtering

Langkah selanjutnya yang diambil oleh peneliti adalah penerapan sistem. Proses ini melibatkan pembuatan kode untuk Sistem Rekomendasi Penjualan Elektronik menggunakan Metode *Content Based Filtering* di Toko Elektronik. Sistem rekomendasi berbasis konten pada dasarnya menggunakan kesamaan atau perbedaan dengan preferensi pengguna untuk menyarankan item yang relevan [3]. Contohnya, kesamaan adalah elemen kunci dari algoritma yang diperlukan untuk merekomendasikan produk elektronik yang dijual, dengan memeriksa kesamaan deskripsi produk yang dapat menjadi rekomendasi produk untuk konsumen, seperti nama barang, tipe, merek, harga, dan diskon barang. Metode ini melibatkan algoritma perhitungan TF-IDF, yang mengukur bobot frekuensi kemunculan sebuah kata dalam data atau Term Frequency (TF) dan Inverse Document Frequency (IDF) [6]. Cosine Similarity juga digunakan untuk menghitung kemiripan antara data pada dataset dengan menganalisis nilai kemiripan melalui bobot vektor setiap data dalam dataset [4]. Tahap ini menghasilkan pengembangan

sistem rekomendasi berbasis konten yang digunakan untuk menganalisis data barang elektronik yang dijual, sehingga memberikan rekomendasi barang di toko elektronik.

4.1. Term Frequency Inverse Document Frequency (TF-IDF)

Term Frequency (TF) dihitung dari frekuensi munculnya kata dari dokumen, dan *Inverse Document Frequency* (IDF) dihitung dari banyaknya istilah dokumen dalam koleksi.

Berikut rumus *Term Frequency* (TF):

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

Berikut rumus *Inverse Document Frequency* (IDF):

$$IDF(t) = \log D/df \quad (2)$$

Berikut rumus *TF-IDF*:

$$Wd, t = TF(d, t) \times IDF(t) \quad (3)$$

4.2. Cosine Similarity

Cosine Similarity menghitung kemiripan sudut anatar vektor dari dua data.

Berikut rumus *Cosine Similarity*:

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (4)$$

Hasil dan Pembahasan

Peneliti menggunakan data barang yang telah didapatkan dari salah satu toko elektronik sebanyak 175 barang elektronik yang dijual ditoko elektronik tersebut. Terdapat 5 variabel yang dibutuhkan untuk memberi rekomendasi barang elektronik pada konsumen yaitu nama barang, merk, tipe, harga, dan diskon barang elektornik.

Data barang yang telah ada dalam excel tersebut digunakan sebagai dataset untuk sistem rekomendasi. Dari data tersebut dilakukan penggabungan data yang kemudian normalisasi data agar dapat diproses oleh sistem.

Data yang telah diproses kemudian dihitung menggunakan rumus TF-IDF. Perhitungan pada rumus *Term Frequency* dapat dilihat pada persamaan (1) untuk menghitung frekuensi term dalam setiap dokumen dan menerapkan transformasi logaritmik yang menghasilkan nilai vektor tiap fitur dari 175 data sebagai berikut:

Tabel 1. Tabel Hasil *Term Frequency*

Fitur	TF
1000000	0.0
1100000	0.0

1110000	0.0
1350000	0.0
1410000	0.0
...	...
Tb Toshiba	0.0
TCL	0.0
TCL 2650000	0.0
Toshiba	0.0

Dalam penghitungan rumus *Inverse Document Frequency* (IDF) yang dapat dilihat dalam persamaan (2) menghitung IDF berdasarkan jumlah dokumen yang mengandung term tertentu dan menerapkan transformasi logaritmik yang mendapatkan hasil penghitungan vektor 175 data sebagai berikut:

Tabel 2. Tabel Hasil *Inverse Document Frequency* (IDF)

Fitur	TF
1000000	4.784190
1100000	4.561046
1110000	5.071872
1350000	4.784190
1410000	5.071872
...	...
Tb Toshiba	5.071872
TCL	4.224574
TCL 2650000	5.071872
Toshiba	4.561046

Setelah melalui penghitungan TF dan IDF maka dihitung vektor tiap fitur data menggunakan rumus TF-IDF yang dapat dilihat dalam persamaan (3) untuk menghitung gabungan dari hasil nilai vektor TF dan IDF sehingga menghasilkan perhitungan TF-IDF fitur dari 175 data berikut:

Tabel 3. Tabel Hasil TF-IDF

Fitur	TF
1000000	0.0
1100000	0.0
1110000	0.0
1350000	0.0
1410000	0.0
...	...
Tb Toshiba	0.0
TCL	0.0
TCL 2650000	0.0
Toshiba	0.0

Tahapan selanjutnya dalam metode *content based filtering* adalah menghitung nilai kemiripan dari produk yang ada pada dataset dengan penghitungan *cosine similarity* yang dapat dilihat rumusnya pada persamaan (4) yang akan menghasilkan nilai kosinus similaritas sebagai berikut:

Tabel 4. Tabel Hasil *Cosine Similarity*

Barang	speaker aktif	...	mesin cuci 1 tb
speaker aktif	1	...	0
speaker aktif	0,81349549 0207682	...	0
speaker aktif	0,50718117 4727819	...	0
...	...	...	...
led	0,05648910 58914419	...	0
led	0,04893776 84828342	...	0
led andorid	0,04358190 93051571	...	0
...	...	...	...
led	0	...	0,20883941 5512876
mesin cuci 2 tb	0	...	0,77185685 8451857
mesin cuci 2 tb	0	...	1

Kemudian peneliti menerapkan sistem rekomendasi *content based filtering* untuk memberikan rekomendasi produk elektronik untuk produk Lemari Es 1 Pintu merk Polytron dengan harga Rp. 2000.000,00 hingga menemukan bahwa ada 2 produk yang memiliki spesifikasi yang dicari yaitu pada Lemari Es 1 Pintu merk Polytron dengan tipe PRA 16 dan PRA 17 seharga Rp. 2000.000,00 produk tersebut kemudian dicari kemiripannya dengan produk lain pada dataset hingga menemukan 4 produk dengan nilai kemiripan terdekat dengan kedua produk yang ditemukan yaitu pada produk berikut:

Tabel 5. Tabel Produk Rekomendasi

Barang	Merk	Tipe	Harga	Similaritas (%)
lemari es 2 pt	polytron	prw 25	4700000	48.989782
lemari es 1 pt	polytron	pra 18	2100000	64.861146
lemari es 1 pt	polytron	prb 159	1900000	62.525402
lemari es 1 pt	polytron	prb 177	2154000	52.549839

Pada hasil rekomendasi produk elektronik menggunakan metode *content based filtering* mendapatkan hasil nilai similaritas yang telah diubah menjadi persen yang kemudian dapat digunakan karyawan untuk memberi rekomendasi produk berdasarkan preferensi dari konsumen. Hasil rekomendasi menunjukan nilai tertinggi dari produk yang paling mirip adalah Lemari Es 1 Pintu merk Polytron tipe PRA 18 harga Rp. 2.100.000,00 dan nilai similaritas atau kemiripan 64.86%.

Jika produk yang diinginkan konsumen tidak ada pada dataset maka sistem tetap akan menampilkan produk yang mirip dengan melihat variabel nama barang, merk dan harganya. Misalnya produk Mesin Cuci 1 Tabung merk Samsung harga Rp. 3.000.000,00 yang tidak ada pada dataset maka akan menampilkan 5 produk dengan nilai maksimum similaritasnya 1.0 diantaranya:

Tabel 6. Tabel Produk Rekomendasi Barang Tidak Ditemukan

Barang	Merk	Tipe	Harga	Max Similaritas
mesin cuci 1 tb	sharp	esg 876	2740000	1.0
mesin cuci 1 tb	sharp	esh 758	3140000	1.0
mesin cuci 1 tb	polytron	paw 7572 7kg	2400000	1.0
mesin cuci 1 tb	polytron	paw 8028 8kg	2850000	1.0
mesin cuci 1 tb	sharp	esh 950	3190000	1.0

Hasil dari perhitungan metode *content based filtering* menghasilkan rekomendasi produk yang ditemukan pada dataset sehingga karyawan dapat memberi rekomendasi produk pada konsumen berdasarkan daftar barang pada hasil rekomendasi dengan nilai tertinggi kemiripannya. Jika produk tidak ada maka akan menampilkan produk yang direkomendasikan berdasarkan nilai maksimum similaritasnya dengan produk yang diinginkan konsumen.

4. Kesimpulan

Berdasarkan hasil implementasi sistem rekomendasi *content based filtering* maka dapat disimpulkan bahwa:

1. Sistem rekomendasi elektronik menggunakan metode *content based filtering* berhasil dibangun menggunakan bahasa pemograman Python, dengan data barang yang ada pada suatu toko elektornik, sehingga dapat membantu karyawan dalam melayani konsumen yang mencari produk berdasarkan preferensi konsumen tersebut.
2. Berdasarkan hasil pengujian sistem dapat berjalan dengan baik. Dari hasil analisis kemiripan, produk yang paling mirip dengan lemari es 1 pintu merk Polytron tipe PRW 25 adalah lemari es 1 pintu merk Polytron tipe PRA 18 dengan harga Rp 2.100.000 dan tingkat kesamaan 64,86%. Produk lain yang juga memiliki tingkat kesamaan tinggi adalah lemari es 1 pintu merk Polytron tipe PRB 159 dengan harga Rp 1.900.000 dan tingkat kesamaan 62,53%, serta lemari es 1 pintu merk Polytron tipe PRB 177 dengan harga Rp 2.154.000 dan tingkat kesamaan 52,55%. Meskipun lemari es

- 2 pintu merk Polytron tipe PRW 25 dengan harga Rp 4.700.000 memiliki tingkat kesamaan terendah yaitu 48,99%, produk ini tetap relevan sebagai alternatif rekomendasi. Maka pengguna dapat memberikan rekomendasi produk yang dicari oleh konsumen. Sistem menampilkan rekomendasi produk dengan nilai kemiripannya paling mendekati produk yang diinginkan konsumen.
3. Berdasarkan hasil pengujian sistem ketika barang yang diinginkan konsumen tidak ditemukan pada dataset maka sistem tetap menampilkan rekomendasi produk yang dapat karyawan informasikan pada konsumen. Dari analisis, seluruh produk mesin cuci 1 TB yang direkomendasikan memiliki tingkat kesamaan maksimal yaitu 1,0. Produk-produk tersebut meliputi mesin cuci merk Sharp tipe ESG 876 dengan harga Rp 2.740.000, tipe ESH 758 dengan harga Rp 3.140.000, dan tipe ESH 950 dengan harga Rp 3.190.000. Selain itu, terdapat juga mesin cuci merk Polytron tipe PAW 7572 7kg dengan harga Rp 2.400.000 dan tipe PAW 8028 8kg dengan harga Rp 2.850.000. Penelitian ini memiliki potensi untuk dikembangkan dengan menghitung presisi sistem guna mengukur tingkat akurasi kemiripan setiap kueri pada dataset yang digunakan [10].

Daftar Rujukan

[1] Batari, F., Larasati, A., & Februariyanti, H. (2021). Sistem Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode Content - Based Filtering. 4(1), 45–54.

[2] Dan, A., & Metode, I. (2020). Collaborative Filtering Untuk Sistem Rekomendasi Konsentrasi Di Stmik Stikom Bali. 9, 171–180.

[3] Fitriani, F. A. (2022). Sistem Rekomendasi Pemilihan Produk Skincare Dengan Pendekatan Contentbased Filtering.

[4] Kurniaji, A., Candra, R., & Santi, N. (2023). Implementasi Metode Content Based Filtering Pada Pemilihan Komik. 10(2).

[5] Susianto, D., & Rusdi, Z. (n.d.). Sistem Rekomendasi Pada Penjualan Elektronik.

[6] Zakiyudin, H., & Marzuki, K. (n.d.). Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta Application of the Cosine Similarity Algorithm and Weighting of the TF-IDF System for New Student Admissions on Private Campuses. 3(1), 19–27. <https://doi.org/10.30812/bite.v3i1.1110>.

[7] Zayyad, M. R. A. (2021). Sistem Rekomendasi Buku Menggunakan Metode Content Based Filtering.

[8] Saputra, F. (2023). Implementasi Algoritma Content Base Filtering Untuk Meningkatkan Efektivitas Penjualan Produk Pada Toko Perlengkapan Sekolah Berbasis Web

[9] Ahmad Syaifudin, M. N. (2024). Penerapan Metode Content-Based Filtering Dalam Strategi Komunikasi Pemasaran Pada Marketplace Tokopedia.

[10] Rio Al Rasyid, D. H. U. N. (2022). Penerapan Algoritma TF-IDF dan Cosine Kemiripan untuk Query Pencarian Pada Dataset Destinasi Wisata.

[11] Ulfa, A., Winarso, D., & Arribe, E. (2020). Sistem Rekomendasi Jurusan Kuliah Bagi Calon Mahasiswa Baru Menggunakan Algoritma C4. 5. Jurnal Fasilkom, 10(1), 61–65.

[12] Sika, S. N. R. (2021). Sistem informasi persediaan stok barang berbasis web pada toko putra gresik. Jurnal Fasilkom, 11(3), 157–16