



## Sistem Rekomendasi Produk Elektronik Berbasis *Semantic Content-Based Filtering* dan *Collaborative Filtering* Menggunakan Pendekatan *Hybrid Weighted*

Ramadhani Rizqi Faisal<sup>1</sup>, Ilham Saifudin<sup>2</sup>, Ginanjar Abdurahman<sup>3</sup>

<sup>1,2,3</sup> Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

[ramadhanirizqifaisal@gmail.com](mailto:ramadhanirizqifaisal@gmail.com), [ilham.saifudin@unmuhjember.ac.id](mailto:ilham.saifudin@unmuhjember.ac.id), [abdurrahmanginajar@gmail.com](mailto:abdurrahmanginajar@gmail.com)

### Abstract

The growth of e-commerce platforms has increased the volume and variety of electronic products, often causing users to struggle to find products that meet their needs—a phenomenon known as information overload. Recommendation Systems help users discover relevant products. However, single methods such as Content-Based Filtering and Collaborative Filtering have limitations, including an inability to grasp Semantic relationships between products, the "cold start" problem, and data sparsity. This study aims to develop an electronic product Recommendation System by combining Semantic Content-Based Filtering and Collaborative Filtering using the Singular Value Decomposition (SVD) algorithm within a Hybrid Weighted framework. The dataset utilized is derived from Amazon Reviews 2023, comprising a total of 90,640 electronic product reviews. The research process involves data preprocessing, generating Semantic representations using Word2Vec, calculating product Similarity via Cosine Similarity, modeling user preferences using SVD, and integrating the two methods using a Hybrid Weighted approach. Evaluation was conducted using Precision, Recall, and Normalized Discounted Cumulative Gain (NDCG) metrics. Results indicate that under the primary configuration—with Alpha = 0.5, an 80:20 data Split, and Top-10 Recommendations—the System achieved a Precision@10 of 0.8692, Recall@10 of 0.0371, and NDCG@10 of 0.8828. Meanwhile, peak performance was observed at Alpha = 0.8, yielding Precision@10 = 0.9230, Recall@10 = 0.0441, and NDCG@10 = 0.9313. These findings demonstrate that the Hybrid Weighted approach generates more relevant electronic product Recommendations and superior ranking quality compared to using single methods.

**Keywords:** Recommender System, Semantic Content-Based Filtering, Collaborative Filtering, Singular Value Decomposition, Hybrid Weighted, Word2Vec.

### Abstrak

Perkembangan platform *e-commerce* menyebabkan jumlah dan variasi produk elektronik semakin meningkat, sehingga pengguna sering mengalami kesulitan dalam menemukan produk yang sesuai dengan kebutuhan mereka (*information overload*). Sistem rekomendasi menjadi salah satu solusi untuk membantu pengguna memperoleh produk yang sesuai. Namun, metode tunggal seperti *content-Based filtering* maupun *collaborative filtering* masih memiliki keterbatasan, seperti ketidakmampuan memahami hubungan semantik antarproduk, *cold start*, dan *data sparsity*. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi produk elektronik dengan menggabungkan *Semantic content-based filtering* dan *collaborative filtering* menggunakan algoritma *Singular Value Decomposition* (SVD) melalui pendekatan *hybrid weighted*. Dataset yang digunakan berasal dari Amazon Reviews 2023 dengan total 90.640 data ulasan produk elektronik. Tahapan penelitian meliputi *preprocessing* data, pembentukan representasi semantik menggunakan *Word2Vec*, perhitungan kemiripan produk menggunakan *cosine similarity*, pemodelan preferensi pengguna menggunakan SVD, serta penggabungan kedua metode menggunakan *hybrid weighted*. Evaluasi dilakukan menggunakan metrik *Precision*, *Recall*, dan *Normalized Discounted Cumulative Gain* (NDCG). Hasil penelitian menunjukkan bahwa pada konfigurasi utama dengan Alpha = 0,5, Data Split 80:20, dan Top-10, sistem memperoleh Precision@10 sebesar 0,8692, Recall@10 sebesar 0,0371, dan NDCG@10 sebesar 0,8828. Sementara itu, performa terbaik diperoleh pada alpha = 0,8 dengan Precision@10 sebesar 0,9230, Recall@10 sebesar 0,0441, dan NDCG@10 sebesar 0,9313. Hasil tersebut menunjukkan bahwa pendekatan *hybrid weighted* mampu menghasilkan rekomendasi produk elektronik yang lebih relevan dan memiliki kualitas pemeringkatan yang lebih baik dibandingkan dengan penggunaan metode tunggal.

**Kata Kunci:** Sistem Rekomendasi, *Semantic Content-Based Filtering*, *Collaborative Filtering*, *Singular Value Decomposition*, *Hybrid Weighted*, *Word2Vec*.

### 1. Pendahuluan

Perkembangan platform *e-commerce* telah meningkatkan jumlah dan keragaman produk elektronik yang tersedia sehingga memberikan lebih banyak alternatif bagi pengguna dalam memilih produk. Kondisi tersebut menimbulkan permasalahan *information overload*, yaitu kesulitan pengguna dalam menemukan produk yang sesuai di antara banyaknya pilihan yang tersedia. Sistem rekomendasi menjadi salah satu solusi yang banyak diterapkan untuk membantu pengguna memperoleh produk yang relevan berdasarkan karakteristik produk maupun

Sistem Rekomendasi Produk Elektronik Berbasis *Semantic Content-Based Filtering* dan *Collaborative Filtering* Menggunakan Pendekatan *Hybrid Weighted*

pola interaksi pengguna, sehingga dapat meningkatkan pengalaman pengguna dalam berbelanja secara daring (Saifudin & Triyanna Widiyaningtyas, 2024). Sistem rekomendasi juga berfungsi sebagai penyaring informasi yang membantu pengguna menemukan produk sesuai dengan preferensi, karakteristik *item*, serta riwayat interaksi pengguna (Roy & Dutta, 2022).

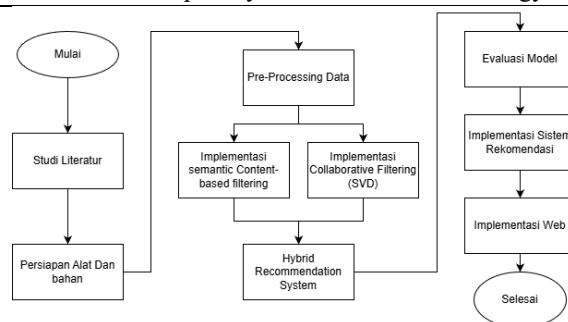
Berbagai penelitian telah mengembangkan sistem rekomendasi menggunakan pendekatan *content-based filtering* maupun *collaborative filtering*. Secara umum, metode sistem rekomendasi dapat dikelompokkan menjadi *content-based filtering*, *collaborative filtering*, dan *hybrid filtering*, yang masing-masing memiliki kelebihan serta keterbatasan (Alfaifi, 2024). *Content-based filtering* memanfaatkan informasi yang terdapat pada atribut produk, seperti nama, kategori, dan deskripsi, sedangkan *collaborative filtering* memanfaatkan pola interaksi pengguna, seperti data *rating*, untuk memprediksi preferensi terhadap produk tertentu. Penelitian terbaru menunjukkan bahwa penggabungan kedua metode melalui pendekatan *hybrid recommendation system* mampu mengurangi kelemahan masing-masing metode tunggal, terutama dalam mengatasi *cold-start* dan data *sparsity*, sehingga menghasilkan rekomendasi yang lebih akurat dan relevan (Rey Muhamad Rifqi et al., 2025). Sistem rekomendasi *hybrid* dapat meningkatkan kualitas rekomendasi karena memanfaatkan kelebihan beberapa metode sekaligus serta mengurangi kelemahan yang terdapat pada metode tunggal (Sabiri et al., 2025). Penggabungan beberapa teknik rekomendasi juga dapat memberikan performa yang lebih baik dibandingkan dengan penerapan masing-masing metode secara terpisah (Pande et al., 2022).

Meskipun demikian, masih terdapat beberapa keterbatasan dalam penelitian sebelumnya. Sebagian besar penelitian *content-based filtering* masih menggunakan representasi berbasis kata kunci atau atribut sederhana sehingga belum mampu menangkap hubungan semantik antarproduk secara lebih mendalam. Pemanfaatan informasi tekstual seperti nama produk, deskripsi, dan ulasan dapat memperkaya representasi *item* sehingga sistem mampu memahami hubungan kemiripan antarproduk dengan lebih baik (Karabila et al., 2023). Di sisi lain, penelitian *collaborative filtering* umumnya hanya memanfaatkan pola interaksi pengguna tanpa mempertimbangkan informasi tekstual produk, sehingga kualitas rekomendasi masih dipengaruhi oleh keterbatasan data interaksi. Keterbatasan jumlah data interaksi dapat menimbulkan permasalahan data *sparsity* dan *cold-start* yang memengaruhi kemampuan sistem dalam menghasilkan rekomendasi untuk pengguna maupun produk baru (Ramakrishna et al., 2023). Selain itu, beberapa penelitian *hybrid recommendation system* hanya menggabungkan hasil rekomendasi tanpa memanfaatkan representasi semantik berbasis *Word2Vec* sebagai dasar pengukuran kemiripan produk maupun mekanisme pembobotan yang seimbang antara kedua metode (Liyana et al., 2025).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem rekomendasi produk elektronik yang mengintegrasikan *semantic content-based filtering* berbasis *Word2Vec* dengan *Collaborative Filtering* menggunakan algoritma *Singular Value Decomposition* (SVD) melalui pendekatan *hybrid weighted*. Representasi semantik digunakan untuk memahami hubungan makna antardeskripsi produk, sedangkan algoritma SVD digunakan untuk mempelajari pola preferensi pengguna berdasarkan data *rating*. Pendekatan *hybrid* berbobot memungkinkan beberapa model rekomendasi dijalankan secara bersamaan, kemudian hasil dari setiap model digabungkan berdasarkan bobot kontribusinya masing-masing. Penggabungan kedua metode diharapkan mampu menghasilkan rekomendasi yang lebih relevan dibandingkan dengan penggunaan metode tunggal (Fernandes et al., 2023). Kinerja sistem dievaluasi menggunakan metrik *Precision*, *Recall*, dan *Normalized Discounted Cumulative Gain* (NDCG) untuk mengetahui efektivitas pendekatan yang diusulkan dalam meningkatkan kualitas rekomendasi produk elektronik (Saifudin & Triyanna Widiyaningtyas, 2024).

## 2. Metode Penelitian

Metode penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun sistem rekomendasi *hybrid*. Pendekatan kuantitatif digunakan karena penelitian mengolah data numerik dan mengukur kinerja sistem secara objektif (Siroj et al., 2024), sedangkan metode eksperimen digunakan untuk menguji dan mengevaluasi performa algoritma yang diterapkan (Akbar et al., 2023). Tahapan penelitian dilakukan secara sistematis mulai dari pengumpulan data, pengolahan data, pembangunan model, hingga evaluasi kinerja sistem. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur penelitian

## 2.1. Dataset

Data yang digunakan berupa 90.640 data produk elektronik yang diperoleh dari Dataset Amazon Reviews 2023 (McAuley Lab). Dataset tersebut mencakup informasi produk elektronik, seperti *ProductId*, *product\_name*, *description*, *categories*, serta data interaksi pengguna berupa *UserId* dan *Rating*. Atribut-atribut tersebut digunakan sebagai dasar dalam penerapan metode *semantic content-based filtering* dan *Collaborative Filtering (SVD)* pada penelitian ini, sedangkan atribut lainnya yang tidak relevan tidak digunakan dalam proses pembangunan sistem rekomendasi. Berikut Tabel 1 yang memuat variabel penelitian beserta penjelasannya.

Tabel 1. Variabel penelitian

| No | ProductId      | product_name                                                         | description                                                                  | Rating | UserId                           |
|----|----------------|----------------------------------------------------------------------|------------------------------------------------------------------------------|--------|----------------------------------|
| 1  | B0BPBJ6SK<br>W | SAMSUNG Galaxy Buds<br>Pro True Wireless                             | Elevate the everyday with epic<br>sound. Samsung Galaxy Buds<br>Pro...       | 4      | AFZUK3MTBIBEDQOPA<br>K3OATUOUKLA |
| 2  | B0BWVDR<br>R6L | Toshiba Canvio Basics<br>1TB Portable External<br>Hard Drive USB 3.0 | Discover one of the easiest<br>ways to free up space on your<br>computer...  | 5      | AFZUK3MTBIBEDQOPA<br>K3OATUOUKLA |
| 3  | B009UNGY<br>HG | Poly Voyager Legend<br>Portable Charge Case                          | The Plantronics Voyager<br>Legend Charge Case is the<br>perfect accessory... | 5      | AFZUK3MTBIBEDQOPA<br>K3OATUOUKLA |
| 4  | B00NH12E<br>HQ | Amazon Basics Apple<br>Certified 30-Pin to USB<br>Charging Cable     | Product Description Amazon<br>Basics Apple Certified 30-Pin<br>to USB...     | 4      | AFZUK3MTBIBEDQOPA<br>K3OATUOUKLA |
| 5  | B0B8YVVF<br>NT | AcuRite Iris<br>Indoor/Outdoor Wireless<br>Weather Station           | Hyperlocal weather information<br>at your fingertips for home,<br>office...  | 1      | AFZUK3MTBIBEDQOPA<br>K3OATUOUKLA |

## 2.2. Preprocessing

Tahap *pre-processing* bertujuan untuk menyiapkan dataset agar bersih, konsisten, dan siap digunakan dalam proses pembentukan sistem rekomendasi. Proses yang dilakukan meliputi pemeriksaan *missing value*, penghapusan data duplikat, seleksi atribut *UserId*, *ProductId*, *product\_name*, *description*, dan *Rating*, serta konversi identitas pengguna ke dalam bentuk numerik (*UserId*). Selanjutnya, dataset dibagi menjadi data latih dan data uji dengan rasio 80:20 untuk proses pelatihan dan evaluasi model.

Selain itu, dilakukan *pre-processing* pada deskripsi produk melalui tahapan *case folding*, *tokenizing*, *stopword removal*, dan *stemming* untuk menghasilkan representasi teks yang lebih konsisten. Hasil *pre-processing* digunakan sebagai masukan dalam pembentukan representasi semantik menggunakan *Word2Vec*, sedangkan atribut *UserId*, *ProductId*, dan *Rating* digunakan untuk membangun model *collaborative filtering* berbasis *Singular Value Decomposition (SVD)*. Tahap ini menghasilkan dataset yang siap digunakan untuk implementasi *Hybrid Weighted*.

## 2.3. Implementasi Semantic Content-Based Filtering

Implementasi *Semantic Content-Based Filtering* dilakukan untuk menghasilkan rekomendasi produk berdasarkan tingkat kemiripan informasi yang dimiliki oleh setiap produk elektronik. Informasi yang digunakan meliputi *product\_name*, *description*, dan *categories* yang telah melalui tahap *pre-processing*. Deskripsi produk kemudian direpresentasikan ke dalam bentuk vektor menggunakan algoritma *Word2Vec*, sehingga setiap produk memiliki representasi numerik yang mampu menggambarkan hubungan semantik antarkata. Setelah representasi vektor terbentuk, tingkat kemiripan antarproduk dihitung menggunakan metode *cosine similarity* berdasarkan Persamaan (1).

$$\text{Sim}(i, j) = \frac{v_i \cdot v_j}{\|v_i\| \times \|v_j\|} \quad (1)$$

DOI: <https://doi.org/10.69693/ijmst.v4i2.12784>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Pada Persamaan (1),  $\text{Sim}(i,j)$  menyatakan tingkat kemiripan antara produk  $i$  dan produk  $j$ . Simbol  $v_i$  merupakan vektor representasi dari produk  $i$ , sedangkan  $v_j$  merupakan vektor representasi dari produk  $j$ . Nilai kemiripan diperoleh dari hasil *dot product* kedua vektor yang kemudian dinormalisasi menggunakan panjang (norma) masing-masing vektor, sehingga menghasilkan nilai kemiripan pada rentang 0 hingga 1.

Nilai kemiripan yang dihasilkan digunakan sebagai *Content-Based Score* (CB Score) untuk menentukan produk yang memiliki karakteristik paling mirip dengan produk acuan. Produk dengan nilai CB Score tertinggi selanjutnya digunakan sebagai salah satu komponen dalam proses *hybrid weighted*.

#### 2.4. Implementasi SVD Collaborative Filtering

Implementasi *collaborative filtering* menggunakan algoritma *Singular Value Decomposition* (SVD) berdasarkan data *UserId*, *ProductId*, dan *Rating* yang telah melalui tahap *pre-processing*. *Singular Value Decomposition* merupakan teknik *matrix factorization* yang digunakan untuk menguraikan matriks interaksi pengguna dan produk menjadi sejumlah faktor laten, sehingga dapat membantu memprediksi preferensi pengguna terhadap produk yang belum pernah diberi *rating*. Model SVD menghasilkan *Collaborative Filtering Score* (CF Score) sebagai prediksi preferensi pengguna terhadap produk, yang selanjutnya digunakan dalam proses *hybrid weighted* (Alabduljabbar, 2023).

#### 2.5. Implementasi Hybrid Weighted

Implementasi *hybrid weighted* dilakukan dengan menggabungkan *Content-Based Score* (CB Score) dan *Collaborative Filtering Score* (CF Score) untuk menghasilkan skor rekomendasi akhir. Penggabungan informasi karakteristik *item* dengan pola pilihan pengguna dapat menghasilkan rekomendasi yang tidak hanya bergantung pada satu jenis informasi saja (Vahidi Farashah et al., 2021). Sebelum dilakukan penggabungan, CF Score dinormalisasi menggunakan *Min-Max Normalization* agar memiliki rentang nilai yang sama dengan CB Score.

$$CF_{norm} = \frac{CF - CF_{min}}{CF_{max} - CF_{min}} \quad (2)$$

Pada Persamaan 2,  $CF_{norm}$  merupakan CF Score yang telah dinormalisasi, sedangkan  $CF_{min}$  dan  $CF_{max}$  masing-masing menyatakan nilai minimum dan maksimum dari seluruh CF Score.

Selanjutnya, CB Score dan CF Score yang telah dinormalisasi digabungkan menggunakan pendekatan *hybrid weighted* dengan parameter pembobotan  $\alpha$  ( $\alpha$ ) untuk menghasilkan *hybrid score*, sebagaimana ditunjukkan pada Persamaan 3.

$$HybridScore = \alpha \times CF_{norm} + (1 - \alpha) \times CBScore \quad (3)$$

Pada Persamaan (3), *Hybrid Score* merupakan skor rekomendasi akhir,  $CF_{norm}$  merupakan skor hasil *collaborative filtering* yang telah dinormalisasi, CB Score merupakan skor hasil *semantic content-based filtering*, sedangkan  $\alpha$  merupakan bobot kontribusi *collaborative filtering*. Produk kemudian diurutkan berdasarkan *hybrid score* tertinggi untuk menghasilkan rekomendasi kepada pengguna.

#### 2.6. Evaluasi

Evaluasi sistem dilakukan menggunakan metrik *Precision*, *Recall*, dan *Normalized Discounted Cumulative Gain* (NDCG). *Precision* digunakan untuk mengukur ketepatan rekomendasi yang diberikan sistem, *Recall* digunakan untuk mengukur kemampuan sistem dalam menemukan seluruh produk yang relevan, sedangkan NDCG digunakan untuk mengevaluasi kualitas urutan (*ranking*) rekomendasi. Seluruh metrik dihitung menggunakan Persamaan (4) hingga Persamaan (6).

$$\text{Precision@K} = \frac{\text{Jumlah item relevan pada Top-K}}{\text{Jumlah item yang direkomendasikan pada Top-K}} \quad (4)$$

*Precision@K* digunakan untuk mengukur tingkat ketepatan rekomendasi yang dihasilkan sistem. Jumlah item relevan pada Top-K merupakan jumlah produk relevan yang berhasil direkomendasikan, sedangkan jumlah item yang direkomendasikan pada Top-K merupakan total produk yang ditampilkan dalam daftar rekomendasi. Nilai *Precision@K* yang lebih tinggi menunjukkan tingkat ketepatan rekomendasi yang semakin baik.

$$\text{Recall@K} = \frac{\text{Jumlah item relevan pada Top-K}}{\text{Jumlah seluruh item relevan}} \quad (5)$$

dengan *Recall@K* digunakan untuk mengukur kemampuan sistem dalam menemukan item relevan. Jumlah item relevan pada Top-K merupakan jumlah produk relevan yang berhasil direkomendasikan, sedangkan jumlah seluruh item relevan merupakan total produk relevan yang seharusnya dapat direkomendasikan kepada pengguna. Nilai *Recall@K* yang lebih tinggi menunjukkan kemampuan sistem yang semakin baik dalam menemukan produk relevan..

|                                                      |     |
|------------------------------------------------------|-----|
| $\text{NDCG@K} = \frac{\text{DCG@K}}{\text{IDCG@K}}$ | (6) |
|------------------------------------------------------|-----|

dengan NDCG@K digunakan untuk mengukur kualitas urutan rekomendasi dengan mempertimbangkan posisi setiap item relevan dalam daftar Top-K. DCG@K merupakan nilai keuntungan kumulatif berdasarkan urutan rekomendasi yang dihasilkan sistem, sedangkan IDCG@K merupakan nilai keuntungan kumulatif ideal apabila

seluruh item relevan ditempatkan pada urutan terbaik. Nilai NDCG@K yang semakin mendekati 1 menunjukkan kualitas pemeringkatan rekomendasi yang semakin baik. Semakin tinggi nilai *Precision*, *Recall*, dan NDCG, semakin baik kualitas rekomendasi yang dihasilkan oleh sistem.

### 3. Hasil dan Pembahasan

Pengujian dilakukan menggunakan dataset Amazon Reviews 2023 dengan skenario pembagian data 80:20 dan 80:10:10. Kinerja sistem dievaluasi menggunakan metrik *Precision*, *Recall*, dan NDCG untuk menganalisis kualitas rekomendasi yang dihasilkan pada setiap skenario pengujian.

#### 3.1. Hasil Pre-processing

Tahap *pre-processing* menghasilkan dataset yang siap digunakan dalam proses pembentukan sistem rekomendasi. Seluruh atribut yang digunakan berhasil dipersiapkan tanpa ditemukan *missing value*, sedangkan identitas pengguna dikonversi ke bentuk numerik untuk membentuk matriks interaksi pengguna–produk. Selain itu, deskripsi produk diproses melalui tahapan *case folding*, *tokenizing*, *stopword removal*, dan *stemming* sehingga menghasilkan representasi teks yang lebih konsisten. Dataset kemudian dibagi ke dalam dua skenario, yaitu 80:20 (*train:test*) dan 80:10:10 (*train:validation:test*), sebagai dasar pengujian performa model pada tahap evaluasi.

#### 3.2. Pelatihan Model *Semantic Content-Based Filtering* dan *Collaborative Filtering*

Model *semantic content-based filtering* dibangun menggunakan atribut *product\_name*, *description*, dan *categories* yang telah melalui tahap *pre-processing*. Representasi semantik produk dibentuk menggunakan *Word2Vec*, kemudian tingkat kemiripan antarproduk dihitung menggunakan *cosine similarity* untuk menghasilkan *Content-Based Score* (CB Score).

Selanjutnya, model *Collaborative Filtering* dibangun menggunakan algoritma *Singular Value Decomposition* (SVD) berdasarkan data *UserId*, *ProductId*, dan *Rating*. Model menghasilkan *Collaborative Filtering Score* (CF Score) yang digunakan sebagai dasar rekomendasi berdasarkan preferensi pengguna.

#### 3.3. Hasil Hybrid Weighted

Pendekatan *hybrid weighted* menggabungkan *Content-Based Score* (CB Score) dan *Collaborative Filtering Score* (CF Score) untuk menghasilkan *hybrid score* sebagai dasar rekomendasi. Setelah proses normalisasi CF Score, seluruh produk diurutkan berdasarkan nilai *hybrid score*, sehingga diperoleh 10 produk elektronik dengan skor tertinggi sebagai rekomendasi akhir. Hasil rekomendasi sistem ditunjukkan pada Gambar 2.

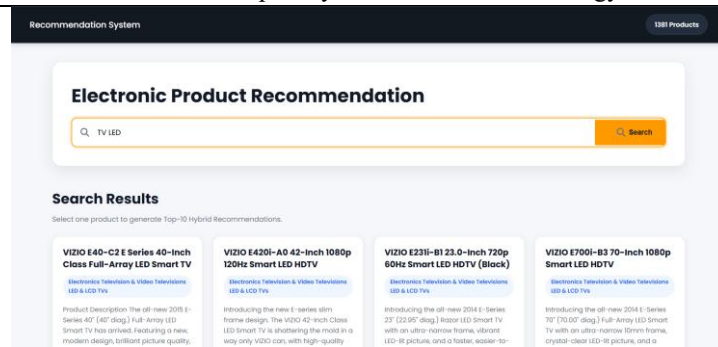
UserId: 4151  
Seed ProductId: B0BBMLD8QT  
Top-5 Hybrid:

|   | ProductId  | Product Name                                       | CB Score | CF Score Normalized | Hybrid Score |
|---|------------|----------------------------------------------------|----------|---------------------|--------------|
| 0 | B01C6ESQM6 | Logitech Z623 400 Watt Home Speaker System, 2....  | 0.8607   | 0.989375            | 0.925037     |
| 1 | B07T99Q9PC | Cyber Acoustics CA-3602FFP 2.1 Speaker Sound S...  | 0.9024   | 0.931000            | 0.916700     |
| 2 | B004VM1T5S | Logitech Wireless Speaker Adapter for Bluetooth... | 0.8750   | 0.941050            | 0.908025     |
| 3 | B09WWG19YS | Bose SoundLink Micro: Small Portable Bluetooth...  | 0.8427   | 0.957225            | 0.899962     |
| 4 | B07P1FN7KQ | Logitech Speaker System Z323 with Subwoofer        | 0.8387   | 0.958825            | 0.898763     |
| 5 | B01MQJ2XE6 | Polk Audio PSW505 12" Powered Subwoofer, 460 W...  | 0.8560   | 0.938825            | 0.897412     |
| 6 | B078WQK4Y5 | JBL GTO Series Premium 6 x 8 Inches Co-Axial S...  | 0.8377   | 0.950675            | 0.894187     |
| 7 | B00GJFGE0K | EIGELIU 3.5 mm Headphone Adapter for Adapter 3...  | 0.8161   | 0.963450            | 0.889775     |
| 8 | B07R8WVY4F | Audio-Technica AT-LP120-USB Direct-Drive Profe...  | 0.8254   | 0.946250            | 0.885825     |
| 9 | B01GR12G9S | Altec Lansing VS2621 2.1 Channel Speaker System    | 0.8799   | 0.890850            | 0.885375     |

Gambar 2. Hasil Top-10 Rekomendasi

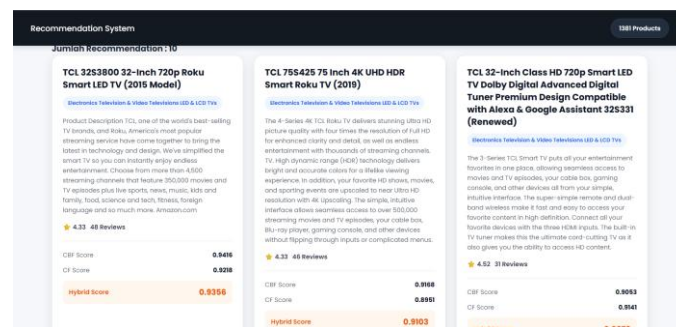
#### 3.4. Implementasi Sistem (*Website*)

Implementasi sistem dilakukan dalam bentuk aplikasi berbasis *web* yang menyediakan fitur pencarian dan rekomendasi produk elektronik dengan pendekatan *hybrid weighted*. Pengguna memilih satu produk dari hasil pencarian sebagai produk acuan (*seed item*), kemudian sistem menghasilkan 10 rekomendasi produk dengan tingkat relevansi tertinggi. Halaman pencarian ditunjukkan pada Gambar 3.



Gambar 3. Tampilan Website Pencarian Produk Elektronik

Gambar 4 menunjukkan sistem berhasil menampilkan 10 rekomendasi produk elektronik berdasarkan produk acuan yang dipilih pengguna. Rekomendasi diurutkan berdasarkan *hybrid score* tertinggi dan disajikan dalam bentuk card yang memuat informasi produk untuk memudahkan pengguna membandingkan hasil rekomendasi. Penerapan sistem rekomendasi *hybrid* pada lingkungan *e-commerce* dapat meningkatkan relevansi rekomendasi karena sistem mempertimbangkan pola interaksi pengguna sekaligus informasi mengenai karakteristik produk (Kwon et al., 2025).



Gambar 4. Tampilan Website Hasil Rekomendasi Hybrid Weighted

### 3.5. Hasil Analisis Parameter Alpha

Tabel 2. Analisis Nilai Alpha

| Alpha      | K         | Precision @10 | Recall@10     | NDCG@10       |
|------------|-----------|---------------|---------------|---------------|
| 0,1        | 10        | 0.6522        | 0.0145        | 0.6692        |
| 0,2        | 10        | 0.7270        | 0.0196        | 0.7433        |
| 0,3        | 10        | 0.7911        | 0.0254        | 0.8057        |
| 0,4        | 10        | 0.8464        | 0.0305        | 0.8592        |
| 0,5        | 10        | 0.8821        | 0.0371        | 0.8918        |
| 0,6        | 10        | 0.9028        | 0.0409        | 0.9105        |
| 0,7        | 10        | 0.9149        | 0.0430        | 0.9232        |
| <b>0,8</b> | <b>10</b> | <b>0.9230</b> | <b>0.0441</b> | <b>0.9313</b> |
| 0,9        | 10        | 0.9208        | 0.0441        | 0.9308        |

Berdasarkan Tabel 2, perubahan nilai *Alpha* berpengaruh terhadap kinerja sistem rekomendasi. Nilai *Precision@10*, *Recall@10*, dan *NDCG@10* cenderung meningkat seiring bertambahnya nilai *Alpha* hingga mencapai performa terbaik pada *Alpha* = 0,8, dengan nilai *Precision@10* sebesar 0,9230, *Recall@10* sebesar 0,0441, dan *NDCG@10* sebesar 0,9313. Hasil tersebut menunjukkan bahwa kontribusi *Collaborative Filtering* yang lebih besar mampu meningkatkan kualitas rekomendasi dalam penelitian ini.

### 3.6. Hasil Analisis Data Split

Tabel 3. Analisis Data Split

| Split    | Precision@10  | Recall@10     | NDCG@10       |
|----------|---------------|---------------|---------------|
| 80:20    | <b>0.8692</b> | 0.0371        | <b>0.8828</b> |
| 80:10:10 | 0.8687        | <b>0.0378</b> | 0.8823        |

Berdasarkan Tabel 3, skenario pembagian data 80:20 menghasilkan nilai *Precision@10* dan *NDCG@10* tertinggi, sedangkan skenario 80:10:10 menghasilkan nilai *Recall@10* yang sedikit lebih tinggi. Secara keseluruhan, kedua skenario menunjukkan performa yang hampir sama, namun pembagian data 80:20 menghasilkan kualitas rekomendasi yang lebih baik.

### 3.7. Hasil Analisis Top-K

Tabel 4. Analisis Top-K

| Algoritma | PRECISION@K (Top-K) |    |    | RECALL@K (Top-K) |    |    | NDCG@K (Top-K) |    |    |
|-----------|---------------------|----|----|------------------|----|----|----------------|----|----|
|           | 5                   | 10 | 15 | 5                | 10 | 15 | 5              | 10 | 15 |

DOI: <https://doi.org/10.69693/ijmst.v4i2.12784>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

|                        |               |        |               |        |        |               |               |        |               |
|------------------------|---------------|--------|---------------|--------|--------|---------------|---------------|--------|---------------|
| <i>SEMANTIC</i>        | <b>0.8914</b> | 0.8622 | 0.8425        | 0.0349 | 0.0623 | <b>0.0845</b> | <b>0.9013</b> | 0.8792 | 0.8650        |
| SVD                    | 0.0022        | 0.0024 | <b>0.0025</b> | 0.0051 | 0.0116 | <b>0.0182</b> | 0.0038        | 0.0063 | <b>0.0084</b> |
| <i>Hybrid Weighted</i> | <b>0.8914</b> | 0.8692 | 0.8519        | 0.0210 | 0.0371 | <b>0.0512</b> | <b>0.9002</b> | 0.8828 | 0.8692        |

Berdasarkan Tabel 4, variasi nilai *Top-K* memengaruhi performa sistem rekomendasi pada setiap algoritma. Metode *hybrid weighted* memberikan *Precision@10* tertinggi sebesar 0,8692, sedangkan metode *Semantic Content-Based Filtering* memperoleh nilai *Recall@15* sebesar 0,0845 dan *NDCG@5* sebesar 0,9013. Sementara itu, metode *Collaborative Filtering* (SVD) menghasilkan nilai evaluasi yang lebih rendah dibandingkan dengan metode lainnya. Hasil tersebut menunjukkan bahwa pendekatan *hybrid weighted* mampu memberikan keseimbangan yang lebih baik antara ketepatan rekomendasi dan kualitas peringkat dibandingkan dengan penggunaan metode tunggal.

### 3.8. Keunggulan dan Keterbatasan Sistem

Sistem yang dikembangkan mampu menggabungkan *semantic content-based filtering* berbasis *Word2Vec* dan *collaborative filtering* menggunakan SVD melalui pendekatan *hybrid weighted* sehingga menghasilkan rekomendasi yang lebih relevan. Selain itu, sistem telah diimplementasikan dalam bentuk aplikasi berbasis *web* dan dilengkapi dengan analisis parameter *alpha*, data *split*, serta *Top-K* untuk mengevaluasi performa model.

Meskipun demikian, penelitian ini masih memiliki keterbatasan, yaitu pengujian hanya dilakukan pada dataset Amazon Reviews 2023 kategori elektronik, nilai *alpha* masih ditentukan secara manual, dan sistem belum terintegrasi secara langsung dengan platform *e-commerce*. Hal tersebut dapat menjadi peluang pengembangan dalam penelitian selanjutnya.

## 4. Kesimpulan

Penelitian ini berhasil mengembangkan sistem rekomendasi produk elektronik menggunakan *semantic content-based filtering*, *collaborative filtering* berbasis *Singular Value Decomposition* (SVD), dan pendekatan *hybrid weighted*. Sistem berhasil diimplementasikan dalam bentuk aplikasi berbasis *web* yang mampu memberikan rekomendasi produk berdasarkan *seed product* yang dipilih pengguna.

Hasil evaluasi menunjukkan bahwa metode *hybrid weighted* menghasilkan performa rekomendasi yang baik dengan nilai *Precision@10* sebesar 0,8692, *Recall@10* sebesar 0,0371, dan *NDCG@10* sebesar 0,8828 pada konfigurasi utama (*Alpha* = 0,5). Performa terbaik diperoleh pada *Alpha* = 0,8 dengan *Precision@10* sebesar 0,9230, *Recall@10* sebesar 0,0441, dan *NDCG@10* sebesar 0,9313, sehingga pendekatan yang diusulkan efektif dalam menghasilkan rekomendasi produk yang relevan.

## Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Bapak Ilham Saifudin, S.Kom., M.Kom., selaku dosen pembimbing yang telah memberikan arahan, bimbingan, dan masukan selama proses penelitian. Ucapan terima kasih juga disampaikan kepada Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember atas dukungan dan fasilitas yang diberikan selama penelitian berlangsung.

## Reference

- Akbar, R., Weriana, W., Siroj, R. A., & Afgani, M. W. (2023). Experimental Research dalam Metodologi Pendidikan. *Jurnal Ilmiah Wahana Pendidikan*, Januari, 9(2), 465–474. <https://doi.org/10.5281/zenodo.7579001>
- Alabduljabbar, R. (2023). *Matrix Factorization Collaborative-Based Recommender System* for Riyadh Restaurants: Leveraging Machine Learning to Enhance Consumer Choice. *Applied Sciences (Switzerland)*, 13(17). <https://doi.org/10.3390/app13179574>
- Alfaifi, Y. H. (2024). Recommender Systems Applications: Data Sources, Features, and Challenges. In *Information (Switzerland)* (Vol. 15, Number 10). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/info15100660>
- Fernandes, L., Miguéis, V., Pereira, I., & e Oliveira, E. (2023). Towards Hyper-Relevance in Marketing: Development of a *Hybrid Cold-start Recommender System*. *Applied Sciences (Switzerland)*, 13(23). <https://doi.org/10.3390/app132312749>
- Karabila, I., Darraz, N., El-Ansari, A., Alami, N., & El Mallahi, M. (2023). Enhancing *Collaborative Filtering-Based Recommender System* Using Sentiment Analysis. *Future Internet*, 15(7). <https://doi.org/10.3390/fi15070235>
- Kwon, Y., Bak, G., & Bae, Y. (2025). A *Hybrid Recommendation System* Based on Similar-Price Content in a Large-Scale *E-commerce* Environment. *Applied Sciences (Switzerland)*, 15(19). <https://doi.org/10.3390/app151910758>
- Liyana, N., Bt, N., Chelvam, M. S., Haw, S.-C., Krisnawati, L. D., & Mahastama, A. (2025). *Hybrid-Based Recommender System* Based on Electronic Product Reviews. *International Journal on Informatics Visualization*. <https://doi.org/10.62527/ijov.9.4.3561>
- Pande, C., Witschel, H. F., & Martin, A. (2022). New *Hybrid Techniques* for Business Recommender Systems. *Applied Sciences (Switzerland)*, 12(10). <https://doi.org/10.3390/app12104804>
- Ramakrishna, M. T., Venkatesan, V. K., Bhardwaj, R., Bhatia, S., Rahmani, M. K. I., Lashari, S. A., & Alabdali, A. M. (2023). *Hybrid Collaborative Filtering* Using Social and Semantic Suggestions for Friend Recommendation. *Electronics (Switzerland)*, 12(6). <https://doi.org/10.3390/electronics12061365>
- Rey Muhamad Rifqi, Djupriadi, Widodo Tri Haryanto, Ema Utami, & Alva Hendi Muhamad. (2025). *Hybrid Recommendation System* in *E-commerce*. *International Journal of Integrated Science and Technology*, 3(5), 1983–1992. <https://doi.org/10.59890/ijist.v3i5.29>
- Roy, D., & Dutta, M. (2022). A Systematic Review and Research Perspective on Recommender Systems. *Journal of Big Data*, 9(1). <https://doi.org/10.1186/s40537-022-00592-5>
- Sabiri, B., Khtira, A., El Asri, B., & Rhanoui, M. (2025). *Hybrid Quality-Based Recommender Systems: A Systematic Literature Review* †. In *Journal of Imaging* (Vol. 11, Number 1). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/jimaging11010012>

- Saifudin, I., & Triyanna Widiyaningtyas. (2024). *Systematic Literature Review on Recommender System: Approach, Problem, Evaluation Techniques, Datasets*. *IEEE Access*, (1). <https://doi.org/10.1109/ACCESS.2024.3359274>
- Siroj, R. A., Afgani, W., Fatimah, F., Septaria, D., & Salsabila, G. Z. (2024). Metode penelitian kuantitatif pendekatan ilmiah untuk analisis data. *Jurnal Review Pendidikan Dan Pengajaran*, 7(3), 11279–11289. <https://doi.org/10.31004/jrpp.v7i3.32467>
- Vahidi Farashah, M., Etebarian, A., Azmi, R., & Ebrahimzadeh Dastjerdi, R. (2021). A *Hybrid Recommender System* Based-on Link Prediction for Movie Baskets Analysis. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00422-0>