

Penerapan Metode *Content-Based Filtering* pada Sistem Rekomendasi

Tegar Ridwansyah, Betty Subartini, Sisilia Sylviani*

Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Padjadjaran, Sumedang 45363, Indonesia

*Corresponding author e-mail: sisilia.sylviani@unpad.ac.id

Article Info

Received April 2024

Accepted April 2024

Published April 2024

Keyword:

Content-based filtering

Recommendation system

Abstract

Content-based filtering is one of the recommendation system methods that provides recommendations based on the similarity of attributes or items preferred by users. Several researchers in various fields have used content-based filtering in their research. This paper describes several studies in various fields using content-based filtering to solve the problems faced in their research. From these studies, the researchers revealed that the use of content-based filtering can produce simple and efficient recommendations and its application can increase the revenue of a business.

1. Pendahuluan

Peradaban dan gaya hidup masyarakat dipengaruhi oleh perkembangan teknologi informasi termasuk internet yang berkembang pesat di abad ke-21. Kehadiran internet membuat komunikasi, interaksi, dan transaksi dalam jarak jauh lebih cepat, mudah, dan murah [1]. Internet sudah menyatukan orang-orang dari berbagai belahan dunia ke sebuah sistem yang besar dan seperti tanpa batas [2]. Pada dunia modern ini, data terus meningkat dan membuat kebutuhan suatu mesin yang dapat menyaring informasi yang ada untuk memprediksi informasi yang sesuai untuk diberikan ke suatu pengguna [3]. Berbagai algoritma dan pendekatan sudah digunakan untuk membuat sistem rekomendasi tetapi orang-orang membutuhkan sistem rekomendasi yang dipersonalisasi yang akan memenuhi kebutuhan khusus pengguna [4]. Dalam waktu kurang dari dua dekade, sistem rekomendasi telah ada di mana-mana di internet, memberikan penawaran produk dan informasi yang dipersonalisasi kepada pengguna [5]. Istilah personalisasi adalah poros pusat dari sistem rekomendasi [6]. Tujuan utama menggunakan personalisasi adalah untuk menghasilkan rekomendasi berdasarkan preferensi dan minat pengguna [7]. Relevansi dan personalisasi rekomendasi dianggap penting untuk meningkatkan pengalaman suatu pengguna, membantu mereka dalam

menemukan informasi yang sesuai dengan minat dan preferensi pribadi [8]. Personalisasi akan memberikan manfaat yang signifikan dalam meningkatkan kepuasan para pengguna [9].

Rekomendasi telah digunakan dalam berbagai masalah, seperti film, lagu, buku, produk, dan lain-lain. Sistem rekomendasi adalah kategori penyaring khusus untuk informasi yang adaptif [10]. Sistem rekomendasi sebagai solusi yang memberikan informasi yang bersifat personal dengan mempelajari penilaian suatu pengguna terhadap item yang pernah dipilihnya [11]. Tantangan yang sulit untuk sistem rekomendasi adalah menghasilkan rekomendasi sesuai dengan selera pengguna [12]. Selain itu, tingginya tingkat keanekaragaman informasi menyebabkan proses pencarian informasi yang relevan menjadi tantangan tersendiri [13]. Dalam proses rekomendasi, proses penyaringan adalah peran yang paling penting [14]. *Collaborative filtering*, *knowledge-based filtering*, *content-based filtering*, dan *social filtering* merupakan metode penyaringan yang sering digunakan [15]. *Collaborative filtering* merupakan salah satu algoritma yang biasa digunakan untuk membuat sistem rekomendasi dan terbukti memberikan hasil rekomendasi yang sangat baik [16]. *Knowledge-based filtering* merupakan sistem rekomendasi yang didasari

oleh *domain knowledge*, yaitu tentang pengaruh fitur dari item terhadap kebutuhan dan preferensi pribadi pengguna [17]. *Knowledge-based filtering* menggunakan pengetahuan tentang pengguna dan item untuk melakukan pendekatan berbasis pengetahuan dalam menghasilkan rekomendasi, dengan alasan tentang item apa yang memenuhi kebutuhan pengguna [18]. *Content-based filtering* merupakan sistem rekomendasi yang didasari dengan korelasi antara antar item. Metode tersebut menggunakan informasi item yang direpresentasikan sebagai atribut untuk dihitung nilai kemiripan antar item [19]. *Social filtering* merupakan sistem rekomendasi yang mementingkan aspek dan komponen pemberi rekomendasi tradisional secara eksplisit yang dirancang dengan menggunakan entitas sosial dan konteks sosial. Salah satu varian penting dari *social filtering* adalah rekomendasi yang diperoleh dengan mengganti lingkungan pengguna dengan model konteks sosial [20]. *Social filtering* menggunakan konten sosial pengguna berdasarkan jejak mereka di jejaring sosial yang berbeda, untuk merekomendasikan informasi yang relevan dengan profil sosial mereka [21]. Tujuan dari sistem rekomendasi adalah untuk mengeliminasi semua informasi yang tidak relevan dan menyajikan informasi yang relevan terhadap suatu pengguna [22]. Tujuan sistem rekomendasi dari perspektif yang luas adalah untuk memberikan rekomendasi yang "baik" dan "berguna" yang membuat pengguna "senang" dengan memuaskan kebutuhan pengguna [23]. Sistem rekomendasi diharapkan mampu memberikan rekomendasi kepada pengguna dengan memanfaatkan informasi yang ada serta memahami preferensi pengguna berdasarkan pilihannya [24].

Salah satu metode yang populer untuk membuat sistem rekomendasi adalah *collaborative filtering* (CF) [25]. Metode tersebut menggunakan informasi preferensi suatu kelompok pengguna untuk dijadikan saran atau prediksi kepada suatu pengguna [26]. Banyak sekali bentuk dari metode CF, tapi terdapat dua langkah utama dari sistem dasar: mendefinisikan kelompok pengguna dengan preferensi yang mirip terhadap suatu pengguna, kemudian pilih item berdasarkan preferensi dari pengguna lain terdekat dan sarankan item tersebut ke suatu pengguna [27]. Tradisional CF menggunakan teknik penyaringan informasi untuk merekomendasikan item berdasarkan riwayat [28]. Algoritma CF selalu memiliki performa lebih baik dalam memprediksi jika terdapat banyak riwayat pengguna dengan item. Sebaliknya, jika CF tidak mempunyai informasi yang cukup maka akan memiliki performa yang kurang baik

[29]. Dengan kata lain, CF tidak bisa memberikan rekomendasi item secara efektif apabila tidak ada atau kurangnya data riwayat. Keterbatasan tersebut dikenal sebagai *cold start problem* [30]. Banyak metode yang telah dibuat untuk menyelesaikan masalah tersebut. Salah satu metode yang digunakan adalah *content-based filtering* [31].

Dalam memberikan rekomendasi item serupa kepada suatu pengguna berdasarkan riwayat item yang disukai, sistem rekomendasi *content-based* membutuhkan suatu item-item yang merepresentasikan pengguna dan keseluruhan item [32]. *Content-based filtering* menggunakan konsep kemiripan konten [33]. Pertama, pengguna dan item direpresentasikan pada ruang fitur yang sama. Kemudian, hitung nilai kemiripan antara pengguna dengan item-item. Pada akhirnya, rekomendasi berhasil dibuat berdasarkan kemiripan suatu pengguna terhadap semua item [34]. Teknik *content-based filtering* yang digunakan merekomendasikan item baru dengan membandingkannya dengan item lain yang terdapat pada profil pengguna [35]. Metode *content-based filtering* memiliki performa yang lebih baik ketika pengguna memiliki banyak aktivitas dalam sistem [36]. Kelebihan sistem rekomendasi dengan *content-based filtering* adalah mempunyai kemampuan untuk merekomendasikan item yang sifatnya baru bagi pengguna [37]. Kelebihan utama dari metode ini adalah kemampuan merekomendasikan item yang tidak terlihat sehingga mampu menyelesaikan keterbatasan *cold start problem* terkait dengan item [38]. Metode *content-based filtering* untuk pembuatan rekomendasi dikatakan layak dan sangat efektif [39].

Beberapa penelitian dalam sistem rekomendasi sudah banyak dilakukan yang menggunakan atau menerapkan *content-based filtering* dalam penelitiannya. Seperti Jain dan rekan-rekannya [6] pada penelitiannya, mereka membuat sistem rekomendasi menggunakan *Singular Value Decomposition* (SVD) untuk mereduksi ribuan kata menjadi dua komponen, kemudian memetakan kata-kata menggunakan *Latent Semantic Analysis* (LSA), dan mencari kemiripannya menggunakan *euclidean distance*. Sementara Nguyen beserta rekannya [38] mengimprovisasi metode *content-based filtering* menggunakan *Jaccard Coefficient Index* untuk mengambil kata-kata pada *features* dan membandingkan kemiripannya menggunakan *soft cosine* (*cosine similarity* yang dimodifikasi). Tommy dan rekan-rekannya [40] mempertimbangkan faktor simpel dan efisien dengan menambahkan algoritma apriori untuk memahami pola

atau perilaku pengguna terhadap suatu item sebelum proses menghitung kemiripan menggunakan *cosine similarity* pada sistem rekomendasinya. Pada sektor pariwisata yang diteliti oleh Muneer dan rekan-rekannya [12] juga mempertimbangkan faktor *budget* dan *interest* pada sistem rekomendasi yang dibuat sebelum menghitung nilai kemiripan menggunakan *cosine similarity*, namun di akhir metode mereka membuat dua buah model sistem rekomendasi untuk mengatasi wisatawan yang pernah atau belum pernah mengunjungi suatu tempat wisata, dengan kata lain untuk mengatasi *cold start problem*.

Penerapan metode sistem rekomendasi pada paper ini akan lebih difokuskan pada salah satu metode rekomendasi *content-based filtering*. Penerapan ini dimaksud untuk membantu memahami penerapan *content-based filtering* pada pembuatan sistem rekomendasi. Hal ini digunakan untuk mencari metode *content-based filtering* dari berbagai pendekatan yang digunakan dan mengatasi keterbatasan yang ada pada metode tersebut, seperti *cold start problem*.

Beberapa penelitian telah melakukan berbagai perubahan untuk meningkatkan performa dan mengatasi keterbatasan pada metode *content-based filtering*. Berbagai pendekatan yang serupa telah digunakan dalam mengembangkan metode tersebut. Tujuan dari penelitian ini adalah untuk mengetahui peranan matematika dan pendekatan yang digunakan dalam metode *content-based filtering*.

2. Metode Penelitian

Penelitian yang disajikan pada artikel ini dilandasi oleh beberapa pertanyaan penelitian sebagai berikut:

- Bagaimanakah penerapan dari *content-based filtering* yang digunakan dalam penelitian beberapa tahun ke belakang dalam menangani banyaknya informasi yang tersedia?
- Bagaimanakah pendekatan yang digunakan penelitian-penelitian tersebut dalam pengembangan *content-based filtering*?
- Bagaimana hasil penelitian yang telah ada sejauh ini tentang *content-based filtering* dapat membantu para pengguna dalam memilih informasi yang relevan dengan kebutuhannya?

Untuk menyelidiki pertanyaan-pertanyaan tersebut, peneliti melakukan studi literatur dari berbagai artikel yang telah terpublikasi dalam rentang tahun 2019-2022. Topik dari artikel-artikel yang dijadikan acuan difokuskan pada penggunaan *content-based filtering* pada sistem rekomendasi.

Adapun metode yang digunakan dalam penelitian ini diuraikan sebagai berikut:

- Tinjauan literatur tentang hasil-hasil penelitian yang telah dilakukan dengan metode *content-based filtering*. Tinjauan ini akan melibatkan tinjauan literatur penelitian tentang penggunaan metode *content-based filtering* untuk membantu memilih informasi pada pengguna. Ini akan mencakup ulasan tentang berbagai jenis pendekatan yang digunakan pada metode tersebut.
- Studi implementasi *content-based filtering* dalam penelitian-penelitian yang telah dilakukan. Studi ini akan melibatkan studi tentang berbagai cara *content-based filtering* yang sudah dilakukan yang menjadi fokus penelitian ini, yang terkait dengan sistem rekomendasi.
- Perbandingan hasil kinerja metode *content-based filtering* yang digunakan dalam penelitian-penelitian tersebut dalam menyelesaikan permasalahan pengguna dalam memilih informasi. Perbandingan ini akan melibatkan pengkajian hasil-hasil yang telah diperoleh dalam penelitian yang menggunakan *content-based filtering* pada berbagai jenis penerapan sistem rekomendasi pada berbagai jenis permasalahan. Ini akan mencakup perbandingan kinerja metode *content-based filtering*.

3. Hasil dan Diskusi

Beberapa tahun ke belakang, penggunaan metode *content-based filtering* sudah mulai banyak diterapkan pada berbagai masalah di luar bidang matematika. Seperti yang telah dijelaskan sebelumnya, *content-based filtering* adalah metode pada sistem rekomendasi yang berfokus pada karakteristik item-item yang akan direkomendasikan berdasarkan preferensi atau minat pengguna.

Tommy beserta rekan dalam penelitiannya menerapkan *content-based filtering* pada aplikasi pemesanan produk sebagai sistem rekomendasi. Penelitian mereka bertujuan untuk mencari metode rekomendasi produk yang mampu mengoptimalkan omset yang diperoleh dalam penjualan dan pemasaran produk dengan mempertimbangkan faktor simpel dan efisien.

Pendekatan yang digunakan pada metode *content-based filtering* dalam penelitian Tommy beserta rekan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk membuat matriks kata menjadi vektor dan *cosine similarity* untuk menghitung nilai kemiripan antar kata pada vektor.

Tabel 1. Kesimpulan dari peneliti yang diteliti

Judul	Penulis	Persoalan	Metode	Hasil
Recommender System dengan Kombinasi Apriori dan Content-Based Filtering Pada Aplikasi [40]	Lukas Tommy, Chandra Kirana, and Vivi Lindawati	Mencari metode rekomendasi produk yang mampu mengoptimalkan omset yang diperoleh dalam penjualan dan pemasaran produk dengan mempertimbangkan faktor simpel dan efisien.	Membuat sistem rekomendasi <i>content-based filtering</i> menggunakan TF-IDF matriks dan <i>cosine similarity</i> serta menambahkan algoritma apriori untuk memahami pola belanja pelanggan.	Peneliti berhasil membuat sistem rekomendasi yang memudahkan pelanggan dalam memilih produk yang dibutuhkan lebih cepat dan efisien, penjualan toko mengalami peningkatan omset, serta akurasi sistem mencapai 73%.
Journal Recommendation System Using Content-Based Filtering [6]	Sonal Jain, Harshita Khangarot and Shivank Singh	Mencari metode yang dapat membantu para peneliti dalam memilih jurnal yang cocok dan tepat sebagai tempat publikasi hasil penelitian mereka.	Mengembangkan metode <i>content-based filtering</i> dengan menggunakan <i>Singular Value Decomposition</i> (SVD) untuk mendekomposisi ribuan kata pada dataset jurnal menjadi dua komponen dan <i>Latent Semantic Analysis</i> (LSA) untuk memetakan kata dan file termasuk maknanya untuk dibandingkan kemiripannya menggunakan <i>euclidean distance</i> .	Menghasilkan top 5 rekomendasi jurnal kepada peneliti berdasarkan jarak <i>euclidean</i> terdekat.
Content-Based Collaborative Filtering using Word Embedding: A Case Study on Movie Recommendation [38]	Luong Vuong Nguyen, Tri-Hai Nguyen, and Jason J. Jung	Mencoba untuk mengimprovisasi metode penelitian sebelumnya untuk mengatasi masalah <i>cold start</i> pada sistem rekomendasi.	Mencoba membuat metode bernama <i>content-based CF</i> yang menggunakan <i>Jaccard Coefficient Index</i> untuk mengekstrak <i>features</i> film, <i>soft cosine</i> untuk membandingkan kemiripan <i>features</i> antar film, dan Word2Vec sebagai <i>embedding model</i> .	Peneliti berhasil membuat metode baru yang memiliki nilai akurasi, <i>precision</i> , <i>recall</i> , dan <i>f1 score</i> yang lebih baik dari peneliti-peneliti sebelumnya.
Tour Spot Recommendation System via Content-Based Filtering [12]	Mishal Muneer, Uzair Rasheed, and Sadia Khalid	Mencari metode yang lebih efektif dari metode penelitian sebelumnya mengenai rekomendasi tempat wisata.	Mengembangkan metode peneliti sebelumnya menggunakan matriks TF-IDF untuk mengubah teks menjadi vektor, <i>cosine similarity</i> untuk membandingkan <i>features</i> antar tempat wisata, dan membuat dua model rekomendasi untuk mengatasi kasus wisatawan pernah dan belum pernah mengunjungi suatu tempat wisata.	Peneliti berhasil mengimprovisasi hasil rekomendasi dalam meningkatkan pengalaman wisatawan karena sistem yang dibuat merekomendasikan tempat wisata sesuai dengan <i>budget</i> dan minat wisatawan.

Term Frequency (TF) merupakan istilah membandingkan banyaknya suatu kata muncul dalam dokumen dengan jumlah kata pada dokumen.

$$TF = \frac{\text{banyaknya suatu kata muncul di dokumen}}{\text{total kata yang ada di dokumen}} \quad (1)$$

Inverse Document Frequency (IDF) merupakan istilah yang mencerminkan proporsi dokumen yang mengandung kata yang dicari.

$$IDF = \log \left(\frac{\text{banyaknya dokumen}}{\text{banyaknya dokumen yang mengandung kata yang dicari}} \right) \quad (2)$$

Rumus untuk mencari nilai TF-IDF cukup mengkalikan hasil TF dengan IDF. Dari bobot kata yang sudah dicari menggunakan TF-IDF, kata-kata yang ada pada kalimat akan dibandingkan menggunakan *cosine similarity*. Misal **a** dan **b** merupakan vektor berisikan kata-kata pada suatu kalimat, maka kemiripannya dapat dihitung menggunakan persamaan (3).

$$\cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}} \quad (3)$$

Selain itu, mereka menambahkan algoritma apriori untuk memahami pola belanja pelanggan. Algoritma apriori adalah salah satu algoritma untuk menjalankan pencarian *frequent itemset* dengan *association rules*. Algoritma ini menggunakan pendekatan *level-wise search*, di mana *k-itemset* dipakai untuk memperoleh *(k+1)-itemset*. Proses ini dilakukan sampai tidak ada lagi kombinasi yang dapat dibentuk.

Cara kerja algoritma apriori adalah sebagai berikut:

- Tetapkan *support* minimum.
- Perulangan 1: kalkulasikan item-item dari *support* (transaksi yang mencakup semua item) dengan memindai basis data bagi *1-itemset*, cek apakah dari *1-itemset* di atas *support* minimum. Jika sudah mencukupi *support* minimum, maka *1-itemset* akan menjadi pola *frequent* tinggi.
- Perulangan 2: untuk memperoleh *2-itemset*, diperlukan penggabungan dari *k-itemset* terdahulu, setelah itu hitung *support* dan cek apakah sudah mencukupi *support* minimum. Jika sudah memenuhi, maka *itemset* yang mencukupi akan dipilih menjadi skema *frequent* tinggi.
- Tentukan angka *k-itemset* dari *support* yang sudah mencukupi *support* minimum dari *k-itemset*.
- Kerjakan kembali proses perulangan hingga tidak ditemukan *k-itemset* yang mencukupi *support* minimum.

Pendekatan yang digunakan pada *content-based filtering* dan penambahan algoritma apriori oleh peneliti ini berhasil membuat sistem rekomendasi yang memudahkan pelanggan dalam memilih produk yang dibutuhkan lebih cepat dan efisien, penjualan toko mengalami peningkatan omset, serta akurasi sistem rekomendasi mencapai 73%.

Peneliti lain, yaitu Jain dan rekan-rekannya melakukan penelitian yang serupa dengan Tommy dan rekan-rekannya, namun terdapat sedikit modifikasi dan pendekatan yang berbeda saat mencari nilai kemiripan kata. Mereka menggunakan *Singular Value Decomposition* (SVD) untuk mereduksi ribuan kata menjadi dua komponen, membuat kombinasi dengan *Latent Semantic Analysis* (LSA), dan menghitung nilai kemiripan kata menggunakan *euclidean distance*. Misal matriks $\mathbf{M}_{m \times n}$ adalah matriks ribuan kata yang ada serta $\mathbf{U}_{m \times m}$, $\mathbf{\Sigma}_{m \times n}$, dan $\mathbf{V}_{n \times n}^T$ merupakan matriks – matriks dekomposisi dari matriks $\mathbf{M}_{m \times n}$. Dari matriks-matriks dekomposisi tersebut akan direduksi menjadi dua komponen dengan memisalkan $k = 2$, diperoleh matriks-matriks dekomposisi yang tereduksi $\mathbf{U}_{m \times 2}$, $\mathbf{\Sigma}_{2 \times 2}$, dan $\mathbf{V}_{2 \times n}^T$. Menurut Jain beserta rekannya, dengan mereduksi jumlah kata menjadi kecil akan mempercepat dan mempermudah dalam menghitung nilai kemiripan yang ada. Misal (x_1, y_1) dan (x_2, y_2) merupakan titik data dari pasangan kata hasil pemetaan menggunakan LSA, untuk menghitung euclidean distance gunakan persamaan (4). Metode dengan pendekatan yang digunakan mereka akan menghasilkan rekomendasi berdasarkan lima *euclidean distance* terdekat.

$$d(x, y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (4)$$

Peneliti lain, yaitu Nguyen beserta rekannya melakukan penelitian dengan metode yang sama, yaitu *content-based filtering*. Pendekatan yang digunakan mereka untuk mencoba mengimprovisasi metode penelitian sebelumnya untuk mengatasi *cold start problem* pada metode tersebut. Pendekatan yang digunakan adalah *Jaccard Coefficient Index* untuk mengekstrak kata dari *features* yang ada dan mencari nilai kemiripannya menggunakan *cosine similarity* yang sudah dimodifikasi menjadi *soft cosine* pada persamaan (5). Peneliti ini berhasil membuat pendekatan baru yang mampu mengatasi keterbatasan metode yang dipakai serta memiliki sistem rekomendasi yang lebih baik dari peneliti-peneliti sebelumnya.

$$\text{Sim}(F_{k,a}, F_{k,b}) = \frac{\sum_{i,j}^N F_{k,a}^i F_{k,b}^j}{\sqrt{\sum_{i,j}^N F_{k,a}^i F_{k,a}^j} \sqrt{\sum_{i,j}^N F_{k,b}^i F_{k,b}^j}} \quad (5)$$

Dengan $F_{k,a}$ dan $F_{k,b}$ adalah set dari *features* item a dan item b , seperti nama, kategori, lokasi.

Peneliti lain, yaitu Muneer dan rekan-rekannya melakukan penelitian yang serupa dengan Nguyen. Muneer beserta rekannya membuat sistem rekomendasi yang mampu mengatasi *cold start problem* dengan cara membuat dua model *content-based filtering* di akhir untuk kasus pengguna yang sudah pernah dan belum pernah mengunjungi suatu tempat wisata. Pendekatan yang dilakukan telah berhasil mengimprovisasi hasil rekomendasi dari keterbatasan yang mungkin terjadi pada metode *content-based filtering*.

4. Kesimpulan

Peran matematika dalam perkembangan ilmu pengetahuan pada era perkembangan teknologi yang pesat ini, sangat dibutuhkan tidak hanya dalam lingkup matematika saja, tetapi pada bidang-bidang lain yang mungkin sebelumnya tidak pernah terbayangkan. Materi yang sudah dipelajari di matematika, baik secara langsung atau konsep, banyak dipakai dalam menyelesaikan masalah-masalah di kehidupan sehari-hari baik itu masalah yang mudah maupun yang lebih sulit.

Salah satu penerapan ilmu atau konsep matematika yang digunakan dalam penyelesaian masalah yang cukup rumit di luar bidang matematika itu sendiri adalah penerapan dalam bidang sistem rekomendasi pada metode *content-based filtering* dengan berbagai pendekatan-pendekatan matematis yang dapat digunakan. Dari penelitian-penelitian yang dikaji pada artikel ini, para peneliti secara tidak langsung memberikan respon yang positif terhadap penggunaan konsep matematika pada pendekatan yang digunakan mulai dari menghitung nilai kepentingan kata menggunakan TF-IDF, mendekomposisi ribuan kata menggunakan SVD, membandingkan kemiripan kata menggunakan *cosine similarity* yang berasal dari rumus trigonometri cosinus, serta penggunaan logika matematika pada algoritma yang digunakan pada metode *content-based filtering*. Dengan menggunakan pendekatan-pendekatan tersebut pada metode *content-based filtering*, peneliti mampu mempelajari dan mengatasi keterbatasan yang terjadi pada metode yang digunakan. Hal tersebut sangat membantu dalam penelitian atau pembelajaran *content-based filtering* ke depannya.

Daftar Pustaka

- H.M, J. 2009. *Sistem Teknologi Informasi: Pendekatan Terintegrasi: Konsep Dasar, Teknologi, Aplikasi, Pengembangan dan Pengelolaan* (3rd ed.). Universitas Gadjah Mada.
- Fahrimal, Y. 2018. Netiquette: Etika Jejaring Sosial Generasi Milenial Dalam Media Sosial. *Jurnal Penelitian Pers dan Komunikasi Pembangunan*, 22(1), 69–78. <https://doi.org/10.46426/jp2kp.v22i1.82>.
- Jaja, V. L., Susanto, B., & Sasongko, L. R. 2020. Penerapan Metode Item-Based Collaborative Filtering Untuk Sistem Rekomendasi Data MovieLens. *d'Cartesian: Jurnal Matematika dan Aplikasi*, 9(2), 78. <https://doi.org/10.35799/dc.9.2.2020.28274>.
- Pratiwi, P. S. 2022. Perancangan Sistem Rekomendasi Berbasis Model Ontologi Untuk Rekomendasi Tempat Magang Mahasiswa. *Syntax Literate: Jurnal Ilmiah Indonesia*, 7(6), 7646–7654. <https://doi.org/https://doi.org/10.36418/syntax-literate.v7i6.7498>.
- Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., & Konstan, J. A. 2014. Exploring the filter bubble. In *Proceedings of the 23rd international conference on World wide web* (pp. 677–686). New York, NY, USA: ACM. <https://doi.org/10.1145/2566486.2568012>.
- Jain, S., Khangarot, H., & Singh, S. 2019. *Journal Recommendation System Using Content-Based Filtering*. Singapore: Springer (pp. 99–108). https://doi.org/10.1007/978-981-13-1280-9_9.
- Rahman, M. M., & Abdullah, N. A. 2018. A Personalized Group-Based Recommendation Approach for Web Search in E-Learning. *IEEE Access*, 6, 34166–34178. <https://doi.org/10.1109/ACCESS.2018.2850376>.
- Tjiang, M. 2023. Analisis Sentimen Pengguna Terhadap Rekomendasi Video YouTube dengan Pendekatan Collaborative Filtering. *Proceeding KONIK (Konferensi Nasional Ilmu Komputer)*, 6, 103–106. Accessed: February 29, 2024, <https://prosiding.konik.id/index.php/konik/article/view/209/133>
- Aldan Nur Zen, M., & Sitanggang, A. S. 2023. Analisis Dampak Sosial Media Dalam Pengembangan Sistem Informasi. *Cerdika: Jurnal Ilmiah Indonesia*, 3(7), 671–682. <https://doi.org/10.59141/cerdika.v3i7.647>.
- Schafer, J. Ben, Konstan, J. A., & Riedl, J. 2001. E-Commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, 5(1/2), 115–153. <https://doi.org/10.1023/A:1009804230409>.
- Wahyudi, I. S. 2018. Big data analytic untuk pembuatan rekomendasi koleksi film personal menggunakan Mlib. Apache Spark. *Berkala Ilmu*

- Perpustakaan dan Informasi*, 14(1), 11. <https://doi.org/10.22146/bip.32208>.
12. Muneer, M., Rasheed, U., Khalid, S., & Ahmad, M. 2022. Tour Spot Recommendation System via Content-Based Filtering. In *2022 16th International Conference on Open Source Systems and Technologies (ICOSST)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICOSST57195.2022.10016820>.
 13. Huda, A. A., & Farida, L. D. 2021. Kajian Sistem Rekomendasi Pada Keanekaragaman Podcast. *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 10(3), 603. <https://doi.org/10.35889/jutisi.v10i3.730>.
 14. Wulandari, I. A., Pahu, G. Y. K. S. S., & Rahayu, P. 2020. Peran Ontologi dalam Pengembangan Sistem Rekomendasi pada Domain Online Learning. *Jurnal Komtika (Komputasi dan Informatika)*, 4(1), 1–8. <https://doi.org/10.31603/komtika.v4i1.3535>.
 15. B.Thorat, P., M. Goudar, R., & Barve, S. 2015. Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System. *International Journal of Computer Applications*, 110(4), 31–36. <https://doi.org/10.5120/19308-0760>.
 16. Laksana, E. A. 2014. Collaborative Filtering dan Aplikasinya. *Jurnal Ilmiah Teknologi Infomasi Terapan (JITTER)*, 1(1), 36–40. <https://doi.org/https://doi.org/10.33197/jitter.vol1.iss1.2014.44>.
 17. Rahmawati, S., Nurjanah, D., & Rismala, R. 2018. Analisis dan Implementasi pendekatan Hybrid untuk Sistem Rekomendasi Pekerjaan dengan Metode Knowledge Based dan Collaborative Filtering. *Indonesian Journal on Computing (Indo-JC)*, 3(2), 11. <https://doi.org/10.21108/INDOJC.2018.3.2.210>.
 18. Ameen, A. 2019. Knowledge based Recommendation System in Semantic Web - A Survey. *International Journal of Computer Applications*, 182(43), 20–25. <https://doi.org/10.5120/ijca2019918538>.
 19. Son, J., & Kim, S. B. 2017. Content-based filtering for recommendation systems using multiattribute networks. *Expert Systems with Applications*, 89, 404–412. <https://doi.org/10.1016/j.eswa.2017.08.008>.
 20. Pazos Arias, J. J., Fernández Vilas, A., & Diaz Redondo, R. P. 2012. *Recommender Systems for the Social Web* (Vol. 32). Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-25694-3>.
 21. Madani, Y., Erritali, M., Bengourram, J., & Sailhan, F. 2019. Social Collaborative Filtering Approach for Recommending Courses in an E-learning Platform. *Procedia Computer Science*, 151, 1164–1169. <https://doi.org/10.1016/j.procs.2019.04.166>.
 22. Sahoo, A. K., Mallik, S., Pradhan, C., Mishra, B. S. P., Barik, R. K., & Das, H. 2019. Intelligence-Based Health Recommendation System Using Big Data Analytics. In *Big Data Analytics for Intelligent Healthcare Management* (pp. 227–246). Elsevier. <https://doi.org/10.1016/B978-0-12-818146-1.00009-X>.
 23. Beel, J., Gipp, B., Langer, S., & Breitingner, C. 2016. Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4), 305–338. <https://doi.org/10.1007/s00799-015-0156-0>.
 24. Zen Munawar, Rustiyana, Yudi Herdiana, & Novianti Indah Putri. 2021. Sistem Rekomendasi Hibrid Menggunakan Algoritma Apriori Mining Asosiasi. *TEMATIK*, 8(1), 84–95. <https://doi.org/10.38204/tematik.v8i1.567>.
 25. Muliadi, K. H., & Lestari, C. C. 2019. Rancang Bangun Sistem Rekomendasi Tempat Makan Menggunakan Algoritma Typicality Based Collaborative Filtering. *Techno.Com*, 18(4), 275–287. <https://doi.org/10.33633/tc.v18i4.2515>.
 26. Arief, A. 2016. Rancang Bangun Sistem Rekomendasi Pariwisata Mobile Advertising Menggunakan Metode Hybrid Filtering Sebagai Pemberdayaan Masyarakat Usaha Kecil Menengah (UKM) di Pulau Ternate. *PROtek: Jurnal Ilmiah Teknik Elektro*, 3(1). <https://doi.org/10.33387/protk.v3i1.38>.
 27. Yang, X., Guo, Y., Liu, Y., & Steck, H. 2014. A survey of collaborative filtering based social recommender systems. *Computer Communications*, 41, 1–10. <https://doi.org/10.1016/j.comcom.2013.06.009>.
 28. Chen, R., Hua, Q., Chang, Y.-S., Wang, B., Zhang, L., & Kong, X. 2018. A Survey of Collaborative Filtering-Based Recommender Systems: From Traditional Methods to Hybrid Methods Based on Social Networks. *IEEE Access*, 6, 64301–64320. <https://doi.org/10.1109/ACCESS.2018.2877208>.
 29. Shi, Y., Larson, M., & Hanjalic, A. 2014. Collaborative Filtering beyond the User-Item Matrix. *ACM Computing Surveys*, 47(1), 1–45. <https://doi.org/10.1145/2556270>.
 30. Ahn, H. J. 2008. A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences*, 178(1), 37–51. <https://doi.org/10.1016/j.ins.2007.07.024>.
 31. Honda, K., Notsu, A., & Ichihashi, H. 2009.

- Collaborative Filtering Based on Sequential Extraction of User-Item Clusters. *Transactions of the Institute of Systems, Control and Information Engineers*, 22(10), 364-370. <https://doi.org/10.5687/iscie.22.364>.
32. Lops, P., de Gemmis, M., & Semeraro, G. 2011. Content-based Recommender Systems: State of the Art and Trends. In *Recommender Systems Handbook* (pp. 73-105). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-85820-3_3.
33. Ayushi, S., & Badri Prasad, V. R. 2018. Cross-Domain Recommendation Model based on Hybrid Approach. *International Journal of Modern Education and Computer Science*, 10(11), 36-42. <https://doi.org/10.5815/ijmecs.2018.11.05>.
34. He, X., He, Z., Song, J., Liu, Z., Jiang, Y.-G., & Chua, T.-S. 2018. NAIS: Neural Attentive Item Similarity Model for Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 30(12), 2354-2366. <https://doi.org/10.1109/TKDE.2018.2831682>.
35. Bezerra, B. L. D., & de A.T. de Carvalho, F. 2004. A symbolic approach for content-based information filtering. *Information Processing Letters*, 92(1), 45-52. <https://doi.org/10.1016/j.ipl.2004.06.003>.
36. Glauber, R., & Loula, A. 2019. Collaborative Filtering vs. Content-Based Filtering: differences and similarities. *arXiv preprint arXiv:1912.08932*. <https://doi.org/10.48550/arXiv.1912.08932>.
37. Rendi, M., Jauhari, J., & Rifai, A. 2016. Pengembangan Sistem Citizen Journalism Berbasis Website dengan Metode Content Based Filtering. *Annual Research Seminar: Computer Science and Information and Communications Technology*, 2(1). Accessed: March 2, 2024, <https://seminar.ilkom.unsri.ac.id/index.php/ars/article/view/878/777>.
38. Nguyen, L. V., Nguyen, T.-H., & Jung, J. J. 2020. Content-Based Collaborative Filtering using Word Embedding. In *Proceedings of the International Conference on Research in Adaptive and Convergent Systems* (pp. 96-100). New York, NY, USA: ACM. <https://doi.org/10.1145/3400286.3418253>.
39. Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. 2017. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, 29-39. <https://doi.org/10.1016/j.eswa.2016.09.040>.
40. Tommy, L., Kirana, C., & Lindawati, V. 2019. Recommender System Dengan Kombinasi Apriori dan Content-Based Filtering Pada Aplikasi Pemesanan Produk. *Jurnal Teknoinfo*, 13(2), 84. <https://doi.org/10.33365/jti.v13i2.299>.