

Sistem Rekomendasi Personalized Music dengan Metode Jenis Gaya

Ida Ayu Tri Sabina Putri^{a1}, I Wayan Supriana^{a2}

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹putri.2208561139@student.unud.ac.id
²wayan.supriana@unud.ac.id

Abstract

The rapid growth of the digital era has led to an increase in online music platforms and music users. However, this abundance of choices has resulted in information overload, making it challenging for users to find their favorite music easily. Thus, the aim of this study was to propose an effective music recommendation method that considers user attributes, music genres, and temporal dynamics. The research utilized a collaborative filtering approach, leveraging user data and music preferences to generate relevant recommendations. Thus, the proposed method aimed to address the issue of information overload and provide more personalized and accurate music recommendations. The results of the recommendation experiments demonstrated the positive effects of integrating these three perspectives. The recommendations generated from this approach were able to assist users in finding their preferred music more conveniently, thereby enhancing user satisfaction with online music platforms.

Keywords: Online Music Platforms, Music Recommendation, Information Overload.

1. Pendahuluan

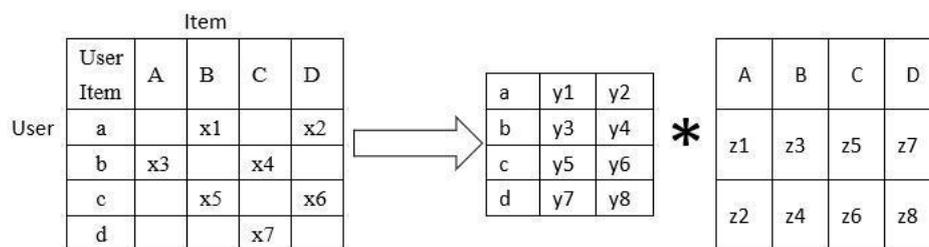
Dengan pertumbuhan terus meningkatnya pengguna jaringan dan perkembangan ekonomi digital, konsumsi digital secara bertahap menjadi kebiasaan dalam konsumsi rumah tangga. Dalam ekonomi digital, konsumsi musik digital menjadi bagian penting, dan hiburan harian masyarakat semakin tergantung pada musik digital, sehingga pasar musik online terus berkembang pesat. Seiring dengan bertambahnya jumlah platform musik dan variasi jenis musik yang tersedia, pengguna memiliki lebih banyak pilihan dan kebebasan dalam memilih, sehingga jumlah pendengar musik online terus meningkat. Namun, seiring dengan pertumbuhan skala musik asli dan peningkatan jumlah serta jenis musik yang pesat, pengguna harus menghabiskan banyak waktu dan energi untuk menemukan musik favorit mereka. Akibatnya, biaya penelusuran pengguna meningkat. Saat mencari musik favorit, orang sering kali harus melalui banyak informasi musik yang tidak relevan, dan dalam proses ini, pengguna rentan menghadapi masalah kelebihan informasi [1]. Jika sistem rekomendasi platform tidak dirancang dengan baik, pengguna akan terus kehilangan minat. Berdasarkan latar belakang tersebut, berbagai metode rekomendasi mulai diterapkan di platform musik. Dalam aplikasi praktis, platform musik utama saat ini melakukan rekomendasi berdasarkan satu metode rekomendasi tunggal. Misalnya, Shrimps Music menggunakan rekomendasi berbasis konten, sementara stasiun musik Last.fm menggunakan rekomendasi penyaringan berbasis sistem [2]. Tingkat akurasi dan cakupan dari kedua jenis metode rekomendasi ini sulit dijamin, dan efek rekomendasi seringkali tidak memenuhi kebutuhan nyata pengguna. Memang, karakteristik pengguna, jenis musik, dan faktor waktu dalam proses rekomendasi seharusnya dipertimbangkan. Namun, pada platform musik saat ini, belum ada penerapan rekomendasi yang mencakup ketiga aspek tersebut secara bersama-sama. Penelitian ini bertujuan untuk mencoba metode rekomendasi baru yang berdasarkan atribut pengguna, jenis musik, dan perpindahan waktu untuk melakukan eksperimen rekomendasi secara Bersama [3], dan diharapkan dapat meningkatkan akurasi dan cakupan rekomendasi platform musik online dengan cara ini.

2. Metode Penelitian

Penelitian ini memperkenalkan algoritma dan metrik rekomendasi yang akan digunakan dalam penelitian ini. Saat ini, filtrasi kolaboratif umumnya dilakukan dengan menggunakan algoritma Top N, sehingga penelitian ini akan menggunakan algoritma tersebut sebagai perhitungan utama. Algoritma rekomendasi linier yang jarang (Sparse Linear Recommendation Algorithm/SLIM) dalam Top N dapat digunakan untuk membangun matriks pengguna melalui pendekatan berbasis model dan memprediksi kesamaan melalui perhitungan berbasis tetangga [4]. Dalam penelitian ini, matriks kesamaan dibangun dan kemudian dilakukan perhitungan kesamaan untuk menghasilkan rekomendasi menggunakan KNN. Bagian berikutnya memperkenalkan algoritma optimisasi peringkat genre musik serta metrik rekomendasi yang akan digunakan dalam eksperimen [5]. Dalam setiap eksperimen sistem rekomendasi, metode rekomendasi merupakan kunci utama, dan penelitian ini menggunakan kombinasi kesamaan pengguna dan jenis lagu untuk menghitung rekomendasi. Langkah pertama adalah pembangunan matriks kesamaan, yang dibangun dengan tujuan untuk mengeksplorasi peringkat pengguna terhadap lagu yang mereka dengarkan pada waktu yang berbeda, sehingga dapat dilihat apakah preferensi mereka berubah seiring perpindahan waktu. Matriks tersebut didekomposisi menggunakan formula (1).

$$R_{m \times n} \approx P_{m \times k} * Q_{k \times n} = \hat{R}_{m \times n} \quad (1)$$

Matriks U dengan ukuran $m \times k$ menggambarkan hubungan antara m pengguna dan k lagu. Matriks V dengan ukuran $k \times n$ menggambarkan hubungan antara k lagu dan n pengguna. Struktur dekomposisi matriks ini ditampilkan dalam Gambar 1.



Gambar 1. Dekomposisi matriks

Dengan merujuk pada matriks kesamaan yang telah dibuat sebelumnya, dilakukan perhitungan kesamaan antar pengguna dengan menggunakan rumus sudut kosinus. Rumus ini digunakan untuk mengukur tingkat kesamaan antara pengguna saat ini dengan daftar lagu yang pernah mereka dengarkan sebelumnya. Dalam perhitungan ini, peringkat pengguna pada waktu yang berbeda saat mendengarkan lagu juga diperhitungkan. Dengan mempertimbangkan atribut pengguna dan perpindahan waktu, rekomendasi dapat disusun dengan lebih baik. Perhitungan kesamaan ini menggunakan rumus utama sebagaimana dinyatakan dalam rumus (2).

$$w_{u,v} = \frac{|N_u \cap N_v|}{\sqrt{|N_u| |N_v|}} \quad (2)$$

- u : mewakili lagu u.
- v : mewakili lagu v.
- N_u : mewakili seberapa disukai lagu u oleh pengguna.
- N_v : mewakili seberapa disukai lagu v oleh pengguna.

Kemudian, dalam penelitian ini digunakan algoritma K Nearest Neighbor (KNN) untuk melatih model rekomendasi. Karena eksperimen ini berdasarkan skor Item untuk membuat prediksi, maka algoritma KNN dipilih sebagai metode pelatihan dalam penelitian ini. Algoritma KNN, yang juga dikenal sebagai algoritma tetangga terdekat K, digunakan untuk menghasilkan rekomendasi dengan menghitung kesamaan antara K pengguna terdekat dengan pengguna target, sesuai dengan rumus (3).

$$d(u, v) = \sqrt{\sum_{k=1}^n |u_k - v_k|} \tag{3}$$

- d : mewakili jarak spasial antara dua pengguna.
- u_k : mewakili K tetangga dari pengguna u.
- v_k : mewakili K tetangga dari pengguna v.

Penelitian ini berfokus pada membuat rekomendasi berdasarkan genre musik dengan algoritma optimasi skor, dan dasar yang paling langsung untuk rekomendasi adalah peringkat pengguna terhadap musik tersebut (Xiang, 2012). Oleh karena itu, eksperimen ini berfokus pada kecenderungan pengguna terhadap genre musik tersebut. Saat ini, genre musik yang paling dominan di platform musik NetEase ditampilkan dalam Tabel 1.

Tabel 1. Genre Musik Dominan Pada NetEase

Kategori Utama	Subkategori
Bahasa	Cina, Eropa dan Amerika, Jepang, Korea, Kanton, Bahasa-bahasa minoritas
Situasi	Pagi, malam, belajar, bekerja, istirahat makan siang, teh sore, naik kereta bawah tanah, mengemudi, olahraga, perjalanan, berjalan, bar
Tema	ACG (Anime, Komik, Game), Children (Anak-anak), Campus (Kampus), Game, Post 70s (Setelah tahun 70-an), Post 80s (Setelah tahun 80-an), Post 90s (Setelah tahun 90-an), Internet Song (Lagu Internet), KTV, Classical (Klasik)
Gaya	Pop, Rock, Folk, Elektronik, Dance, Rap, Musik Ringan, Jazz, Country, R&B/Soul, Klasik, Etnik, Inggris, Metal, Punk, Blues, Reggae, Musik Dunia, Latin, Alternatif/Indie, New Age, Kuno, Post-rock, Bossa Nova
Emosi	Nostalgia, segar, romantis, seksi, sentimental, penyembuhan, relaksasi, kesepian, emosional, bersemangat, bahagia, tenang, nostalgia

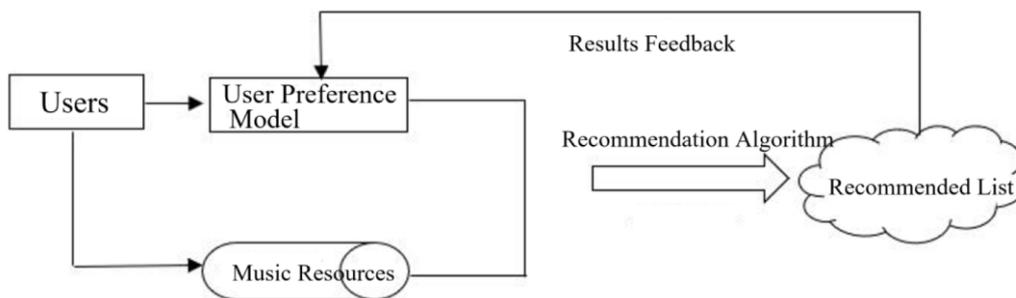
Data dalam penelitian ini utamanya berasal dari platform musik NetEase cloud, yang memiliki jumlah pendengar yang lebih besar dan informasi lagu yang lebih lengkap. Program crawler yang ditulis dalam bahasa Python digunakan untuk mengumpulkan 66.291 informasi data dari platform musik NetEase cloud. Bidang utama dari data yang dikumpulkan dalam penelitian ini adalah id daftar lagu, genre musik dari daftar lagu, jumlah kali lagu diputar, artis, waktu rilis lagu, durasi lagu, dan bidang-bidang terkait lainnya.

Data dalam penelitian ini masih perlu dalam format sistem rekomendasi, sehingga untuk memudahkan partisipasi dalam perhitungan dan menyesuaikan dengan kebutuhan library Python surprise saat ini dan kerangka kerja sistem rekomendasi yang paling umum, format data diproses menjadi format dasar dari dataset movielens. Data yang telah diproses ditampilkan dalam Gambar 4.

Website	id	01_kind	02_kind	create_time	favorite_count	share_count	comment_count	play_count	playlist_count
http://mus	375661355	Subject	After 90	2016/5/5	378664	8285	5243	22238996	178
http://mus	378324005	Language	Chinese	2016/5/9	439451	6161	1441	21703896	210
http://mus	554473977	Subject	List	2017/1/1	283964	2592	2955	20803614	100
http://mus	498708023	Language	Chinese	2016/10/30	464891	6006	2561	20403924	200
http://mus	440561933	Language	Europe and America	2016/8/10	546976	6139	2650	20364706	100
http://mus	758106745	Emotion	Nostalgia	2017/6/10	203546	1171	959	19682608	300
http://mus	497639728	Language	Video Soundtrack	2016/10/29	262802	1485	718	19142126	50
http://mus	375779419	Subject	Excitement	2016/5/5	500907	18925	3743	18364244	59
http://mus	331398396	Emotion	Nostalgia	2016/4/6	500601	5323	1921	17436642	92
http://mus	417794815	Emotion	Chinese	2016/7/8	239929	2131	995	17351472	101
http://mus	330634870	Language	Europe and America	2016/4/5	475196	3379	2642	17258234	181
http://mus	504347934	Language	Chinese	2016/11/7	462340	3291	1241	17223906	53
http://mus	374012053	Language	Europe and America	2016/5/2	267265	1901	1060	16401452	104
http://mus	427850291	Style	Rock and Roll	2016/7/22	478410	3533	2861	16243550	55
http://mus	818563810	Scene	Evening	2017/7/21	337444	11287	22277	15164669	41
http://mus	171664498	Style	Electronic	2016/2/18	683050	13507	3216	15085841	76
http://mus	153864251	Style	Electronic	2016/1/14	313006	2957	1646	14168671	344
http://mus	308905952	Subject	Video Soundtrack	2016/2/28	394877	4910	2209	14157364	100
http://mus	399723525	Language	Europe and America	2016/6/12	433802	6576	2366	14107973	67
http://mus	458329079	Language	Europe and America	2016/9/5	393350	2902	1276	14090920	15
http://mus	517438933	Emotion	Sentimental	2016/11/23	348668	3017	829	13973394	78
http://mus	698720887	Language	Europe and America	2017/4/16	178391	866	517	13914202	101
http://mus	367649303	Emotion	Alone	2016/4/22	341332	2477	1888	13870706	117
http://mus	307592533	Language	Europe and America	2016/2/25	191853	1886	2091	13323389	44
http://mus	449559997	Style	World Music	2016/8/23	354523	3601	2423	13312312	76

Gambar 3. Contoh potret data sebagian

Id menggambarkan nilai id dari daftar lagu; kind menggambarkan jenis gaya dari daftar lagu, yang digunakan untuk perhitungan penilaian gaya kemudian; create_time, yang mewakili waktu pembuatan, digunakan untuk perhitungan migrasi waktu kemudian; dan data pada kolom terakhir adalah jumlah koleksi, jumlah berbagi, jumlah komentar, jumlah pemutaran, dan jumlah lagu, yang akan digunakan dalam penelitian lebih lanjut di masa mendatang.



Gambar 4. Pemodelan

Sistem rekomendasi musik online mirip dengan model sistem rekomendasi tradisional yang terdiri dari tiga komponen utama: modul informasi pengguna, model sumber daya musik (pemodelan objek rekomendasi), dan algoritma rekomendasi musik personal. Hal ini dapat dilihat pada Gambar 4. Fokus dari eksperimen penelitian ini adalah pada modul sumber daya musik, yang terutama berisi beberapa informasi tentang daftar lagu seperti ID daftar lagu, jumlah lagu, penyanyi, dan gaya musik. Rekomendasi utamanya didasarkan pada informasi gaya musik di dalamnya. Pada saat yang sama, dari tiga metode sebelumnya, dapat dilihat bahwa rekomendasi berbasis konten memiliki data yang terbatas di domain musik dan sulit untuk menemukan preferensi baru pengguna. Rekomendasi berbasis aturan asosiasi mempertimbangkan objek yang lebih sedikit dan cakupannya lebih rendah. Rekomendasi musik berbasis jaringan sosial membutuhkan komputasi yang besar dan struktur yang kompleks serta seringkali memerlukan integrasi data lintas platform. Oleh karena itu, masing-masing dari ketiga metode tersebut memiliki kelebihan dan kekurangannya sendiri. Dalam penelitian ini, algoritma rekomendasi berbasis filtrasi kolaboratif digunakan untuk menghindari kekurangan ketiga metode tersebut dengan mempertimbangkan atribut pengguna, jenis gaya, dan migrasi waktu. Selain itu, algoritma ini memiliki volume komputasi yang relatif kecil dan digunakan di berbagai platform musik utama saat ini.

3. Hasil dan Diskusi

Ada tiga metode eksperimen umum yang sering digunakan dalam sistem rekomendasi, yaitu Studi Pengguna, Eksperimen Offline, dan Eksperimen Online. Dalam penelitian ini, metode eksperimen yang digunakan adalah eksperimen offline. Dataset yang ada dibagi menjadi dua bagian berdasarkan perbandingan 8:2, yaitu Training Set dan Test Set. Algoritma rekomendasi berbasis filtrasi kolaboratif terdiri dari dua jenis, yaitu algoritma berbasis pengguna (UserCF) dan algoritma berbasis produk (ItemCF). Pada awal eksperimen, sebagian kecil data digunakan untuk membandingkan kedua algoritma tersebut. Penelitian ini menggunakan nilai $K = 10$, dan hasilnya ditampilkan dalam Tabel 2.

Tabel 2. UserCF and ItemCF

	UserCF	ItemCF
Recall	0.431348	0.43487
Precision	0.37293	0.40143
Popularity	3.45781	3.27159
Coverage	0.51246	0.53841

Berdasarkan tabel di atas, terlihat bahwa indikator-alat dari algoritma filtrasi kolaboratif berbasis pengguna (UserCF), seperti Recall, Precision, Popularity, dan Coverage, memiliki nilai yang lebih rendah dibandingkan dengan algoritma filtrasi kolaboratif berbasis produk (ItemCF). Oleh karena itu, dalam penelitian ini, algoritma filtrasi kolaboratif berbasis produk dipilih. Dalam penelitian ini, dilakukan percobaan dengan variasi nilai K pada algoritma ItemCF, yaitu $K=5, 10, 20$, dan 30 . Hasil percobaan tersebut kemudian ditampilkan dalam Tabel 3.

Tabel 3. ItemCF dengan nilai K yang berbeda

K	5	10	20	30
Recall	0.27364	0.30142	0.31921	0.39281
Precision	3.10323	0.32157	0.38805	3.11451
Popularity	0.51348	0.45841	0.35421	0.51967
Coverage	0.30142	0.31921	0.39281	0.45841

Berdasarkan tabel di atas, terlihat bahwa terdapat pola peningkatan dan penurunan tingkat recall dan akurasi seiring dengan peningkatan nilai K . Titik tertinggi untuk tingkat recall dan akurasi terjadi saat K bernilai 20 , dengan nilai masing-masing sebesar 0.45841 dan 0.38805 . Oleh karena itu, dalam penelitian ini dapat disimpulkan bahwa algoritma lebih akurat saat K bernilai 20 . Penelitian ini melakukan klasifikasi terlebih dahulu terhadap jenis lagu, kemudian melakukan perhitungan peringkat pengguna untuk jenis lagu yang berbeda guna memberikan rekomendasi. Namun, untuk mengkaji pengaruh nilai K yang berbeda, dilakukan perhitungan terpisah menggunakan algoritma UserCF dengan nilai $K 5, 10, 20$, dan 30 . Hasil percobaan ini ditampilkan dalam Tabel 4.

Tabel 4. UserCF dengan nilai K yang berbeda

K	5	10	20	30
Recall	0.31121	0.33429	0.34957	0.34281
Precision	0.15481	0.16249	0.27543	0.16329
Popularity	3.09541	3.11428	3.13027	3.14438

K	5	10	20	30
Coverage	0.509128	0.51459	0.50129	0.49499

Berdasarkan perbandingan eksperimental, dapat dilihat bahwa ketika nilai K adalah 5, 10, 20, dan 30, kedua tabel menunjukkan pola yang sama. Hal ini menunjukkan bahwa ItemCF dan UserCF memiliki sifat yang serupa ketika nilai K mengambil keempat nilai tersebut. Selain itu, pada saat K=20, Recall dan Precision sistem lebih tinggi, menandakan bahwa nilai K tersebut adalah yang paling sesuai. Dengan demikian, eksperimen menunjukkan bahwa ketika nilai K diambil dengan tepat, akurasi dan cakupan rekomendasi dari algoritma rekomendasi berbasis filtrasi kolaboratif dengan klasifikasi gaya musik lebih tinggi. Dengan kata lain, ketika kita menggunakan nilai K yang tepat, tingkat ketepatan dan ketersediaan rekomendasi sistem meningkat.

4. Kesimpulan

Dapat disimpulkan bahwa penelitian ini menemukan bahwa efek rekomendasi dapat dioptimalkan dengan menyesuaikan nilai K menggunakan algoritma KNN. Selain itu, penelitian ini menemukan bahwa pengaturan algoritma rekomendasi gabungan dapat mengoptimalkan sistem rekomendasi hingga batas tertentu. Rekomendasi gabungan berdasarkan perspektif-perspektif seperti atribut pengguna, jenis gaya, dan migrasi waktu dapat mengkompensasi kekurangan metode rekomendasi lainnya dan meningkatkan akurasi. Dalam sistem rekomendasi musik online, algoritma penilaian jenis gaya digunakan untuk menghitung skor tertinggi pada lagu-lagu yang direkomendasikan, dan tiga perspektif yaitu atribut pengguna, gaya musik, dan migrasi waktu dianalisis, dan sistem rekomendasi dibangun dengan menggabungkan ketiga perspektif. Meskipun demikian, penelitian ini juga memiliki beberapa kelemahan. Misalnya, kecerdasan buatan seperti deep learning telah mulai diterapkan dalam bidang sistem rekomendasi, namun penelitian ini tidak mempertimbangkan dan menguji teknologi tersebut. Hal ini juga menjadi arah pengembangan masa depan dari sistem rekomendasi musik online.

Daftar Pustaka

- [1] L. Chen, Y. Yang, N. Wang, K. Yang, and Q. Yuan, "How Serendipity Improves User Satisfaction with Recommendations? A Large-Scale User Evaluation," in *The World Wide Web Conference*, in WWW '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 240–250. doi: 10.1145/3308558.3313469.
- [2] Y. Cheng and X. Bu, "Research on Key Technologies of Personalized Education Resource Recommendation System Based on Big Data Environment," *J. Phys. Conf. Ser.*, vol. 1437, no. 1, 2020, doi: 10.1088/1742-6596/1437/1/012024.
- [3] S. Gopal Krishna Patro et al., "A Conscious Cross-Breed Recommendation Approach Confining Cold-Start in Electronic Commerce Systems," *IEEE Access*, vol. 11, no. August, pp. 82857–82870, 2023, doi: 10.1109/ACCESS.2023.3274844.
- [4] F. Abbas, "One Size Does Not Fit All: Modeling Users' Personal Curiosity in Recommender Systems." [Online]. Available: <https://arxiv.org/abs/1907.00119>
- [5] E. J. Graham et al., "Optimization of a combined power plant CO₂ capture and direct air capture concept for flexible power plant operation," 2024, doi: 10.1039/d4ee00309h.