Repeater : Publikasi Teknik Informatika dan Jaringan
Vol. 2 No. 1 Januari 2024
E-ISSN : 3046-7276, dan P-ISSN : 3046-7284, Hal. 57-65
DOI: https://doi.org/10.62951/repeater.v2i3.140
Available online at: https://journal.aptii.or.id/index.php/Repeater

Implementasi Item-Based Collaborative Filtering Untuk Rekomendasi Film

Rayhan Rizal Mahendra[1], Fetty Tri Anggraeny[2], Henni Endah Wahanani[3]
[1-3]program Studi Informatika, Universitas Pembangunan Nasional "Veteran" Jawa Timur,Indonesia

Alamat: Jl. Rungkut Madya No.1, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur
Email: 20081010045@student.upnjatim.ac.id[1*], fettyanggraeny.if@upnjatim.ac.id[2], henniendah@upnjatim.ac.id[3]

Abstract*.Item-based collaborative filtering is a popular technique in recommendation systems that aims to provide suggestions for films to watch or services to users based onsimilarities between items. In this approach, the similarity between items is calculated using metrics such as cosine similarity, allowing the prediction of user preferences for items that have never been rated. This research implements Item-based collaborative filtering using datasets from Kaggle. Experimental results show that the resulting model is able to provide recommendations with significant improvements in evaluation metricssuch as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) of 3.05 and 3.26. This shows that the smaller the value, the better.*

Keywords: *Item Based Collaborative Filtering, Recommendations, Film Analysis*

Abstrak.Item-based collaborative filtering merupakan salah satu teknik populer dalam sistem rekomendasi yang bertujuan untuk memberikan saran film yang ditonton atau layanan kepada pengguna berdasarkan kesamaan antar item. Dalam pendekatan ini, kesamaan antara item dihitung menggunakan metrik seperti cosine similarity, memungkinkan prediksi preferensi pengguna terhadap item yang belum pernah dinilai. Penelitian ini mengimplementasikan Item-based collaborative filtering dengan menggunakan dataset dari kaggle. Hasil penelitian menunjukkan bahwa model yang dihasilkan mampu memberikan rekomendasi dengan peningkatan signifikan pada metrik evaluasi seperti Mean Absolute Error (MAE) dan Root Mean Square Error(RMSE) sebesar 3.05 dan 3.26. Hal ini menunjukan bahwa semakin kecil nilai maka semakin baik.

Kata kunci: Pemfilteran Kolaboratif Berbasis Item, Rekomendasi, Analisis Film

1. LATAR BELAKANG

Di tengah revolusi teknologi informasi yang sedang berlangsung, Sebagian bahkan seluruh masyarakat akan melihat bahwa teknologi yang sedang berkembang memiliki dampak positif yang membantu dalam pekerjaan di bidang bisnis maupun aktivitas sehari-hari. Di dalam perkembangan teknologi ini tidak lupa juga dengan kecanggihan teknologi yaitu Machine Learning. Salah satu aspek yang paling mencolokdalam evolusi teknologi adalah kemajuan dalam bidang kecerdasan buatan dan machinelearning. Machine learning fokus pada pengembangan sistem yang dapat belajar dan berkembang dari data, tanpa harus diprogram untuk melakukan suatu tugas tertentu. Machine Learning telah melampaui batas-batas yang sebelumnya dianggap tidak mungkin. Salah satu perkembangan dari machine learning ini adalah rekomendasi untuk film. Sistem rekomendasi telah menjadi bagian penting dalam industri hiburan, terutamadalam platform yang menyediakan konten film dan serial televisi.

Dengan meningkatnya jumlah konten yang tersedia, pengguna seringkali mengalami

Received: Juni 12, 2024; Revised: Juni 26, 2024; Accepted: Juli 16, 2024; Published: Juli 18, 2024;

kesulitan dalam menemukan film yang sesuai dengan preferensi mereka. Sistem rekomendasi hadir untuk mengatasi masalah ini dengan memberikan saran yang dipersonalisasi, membantu pengguna menemukan konten yang mungkin mereka sukai berdasarkan perilaku dan preferensi sebelumnya. Dengan kemampuan untuk belajar daridata, mengidentifikasi pola, dan membuat keputusan tanpa intervensi manusia, Machine

Learning menjadi pusat dari revolusi kecerdasan buatan. Sebuah sistem rekomendasi menggunakan filter kolaboratif akan memberikan saran terbaik di pasar[1].

Collaborative Filtering (CF) adalah teknik sistem rekomendasi yang banyak digunakan dan mungkin merupakan yang paling umum. Collaborative filtering memanfaatkan informasi rating dari beberapa pengguna untuk memprediksi rating item untuk pengguna tertentu [2]. Collaborative filtering dibagi menjadi dua macam yaitu user-based collaborative filtering dan item-based collaborative filtering [3]. Peningkatankinerja komputasi, ketersediaan data yang melimpah, dan pemahaman yang mendalam tentang algoritma telah membentuk fondasi bagi perkembangan teknologi. e-commerce memiliki peran penting dalam perekonomian karena dapat menciptakan lapangan kerja, meningkatkan pendapatan masyarakat, serta mendukung pertumbuhan ekonomi local. Sebuah situs web dapat dikenali sebagai kumpulan item-item yang menarik (Xiaoyuan Su, and Taghi M. Khoshgoftaar, 2009)[4].

Metode yang digunakan pada penelitian ini adalah item based collaborative filtering. Tujuan dari penelitian ini adalah untuk mengimplementasikan item based collaborative filtering untuk rekomendasi item.Untuk kontribusi pada penelitian ini adalah mengimplementasikan item based collaborative filtering ke dalam dataset untuk rekomendasi produk.

2 KAJIAN PUSTAKA

2.1 Sistem rekomendasi

Sistem rekomendasi adalah merupakan sistem atau aplikasi yang dibuat untuk dapat menyediakan dan memberikan rekomendasi dari suatu item untuk membuat suatu keputusan yang diinginkan oleh pengguna sistem (Ungkawa et al., 2011)[5]. Sistem ini sangat umum digunakan dalam berbagai platform online seperti e-commerce, layanan streaming, dan media sosial. Sistem rekomendasi terdapat beberapa jenis yaitu content based filtering, collaborative filtering dan hybrid filtering. pada penelitian inimenggunakan collaborative filtering

2.2 Collaborative filtering

Collaborative Filtering adalah salah satu teknik utama dalam sistem rekomendasi yang didasarkan pada analisis pola perilaku pengguna. Tujuan utama dari collaborative filtering

adalah untuk memprediksi preferensi pengguna terhadap item tertentu (misalnya produk, film, atau lagu) berdasarkan kesamaan dengan pengguna lainatau kesamaan antar item. Teknik ini bekerja dengan asumsi bahwa jika dua pengguna memiliki kesamaan dalam menilai beberapa item, maka mereka mungkin akan memilikikesamaan dalam menilai item lainnya.

Ada dua jenis utama dari collaborative filtering yaitu User-Based Collaborative Filtering dan Item-Based Collaborative Filtering. User-Based Collaborative Filtering mencari pengguna lain yang memiliki preferensi serupa dengan pengguna target dan merekomendasikan item yang disukai oleh pengguna-pengguna serupa tersebut.Sebaliknya, Item-Based Collaborative Filtering mencari item yang mirip dengan item yang disukai oleh pengguna target dan merekomendasikan item tersebut. untuk penelitian yang digunakan adalah fokus pada implementasi Item Based Collaborative Filtering.

2.3 Item based collaborative filtering

Item-based collaborative filtering merupakan kebalikan dari metode user- based.Bila pada user-based collaborative filtering prediksi rekomendasi didasarkan oleh kemiripan antar pengguna, pada item- based prediksi untuk rekomendasi didasarkan oleh kemiripan antar item. Adapun cara kerja di dalam metode item-based collaborative filtering antara lain:

2.3.1 Mengidentifikasi Kesamaan Item

Menghitung kemiripan antara item berdasarkan rating yang diberikan oleh pengguna. Metode yang digunakan bisa sama dengan yang digunakan dalam user-based filtering.

2.3.2 Membangun Matriks Kemiripan Item

Membuat matriks yang menunjukkan seberapa mirip satu item dengan itemlainnya.

2.3.3 Memberikan Rekomendasi

Rekomendasi diberikan dengan melihat item yang mirip dengan item yangsudah dinilai tinggi oleh pengguna.

2.3.4 Cosine similarity

*Cosine similarit*y adalah metode yang sering digunakan dalam sistem rekomendasi untuk mengukur seberapa mirip dua item. Dalam konteks item-basedcollaborative filtering, metode ini membantu merekomendasikan item kepada pengguna berdasarkan kesamaan dengan item lain yang sudah diberi rating. Berikut merupakan persamaan untuk mencari kemiripan pada rekomendasi produk:

$$Sim(k,l) = \frac{\sum_{u=1}^{m}(Ru,k-\bar{R}k)(Ru,l-\bar{R}l)}{\sqrt{\sum_{u=1}^{m}(Ru,k-\bar{R}k)(Ru,l-\bar{R}l)^2}\sqrt{\sum_{u=1}^{m}(Ru,k-\bar{R}k)(Ru,l-\bar{R}l)^2}}$$

Untuk menghitung kemiripan item antara (k,l) dengan item yang diberikan oleh masing-masing user, kemudian menjumlahkan R sebagai rating dari item k kemudian dikurangi rata-rata item dari k. kemudian untuk item l, menjumlahkan R sebagai rating dari item l dan dikurangi rata-rata dari item l. kemudian dikalikan antara item k dan l untuk mencari pembilangnya. selanjutnya mencari penyebut dari masing-masing item (k,l) kemudian menggunakan cara yang sama dengan mencari pembilang tetapi setelah itu dikuadratkan kemudian diakarkan. Langkah terakhir membagi antara pembilang dan penyebutnya. hasil yang didapatkan berkisar antara -1, 0 dan 1. jika kurang dari -1 makanilai kemiripan sangat berbanding terbalik. jika nilai kemiripan 0 maka kemungkinan ada nilai kemiripan antara item, dan jika nilai 1 maka, nilai kemiripan item sangat tinggi.

2.3.5 Prediksi

Item-based collaborative filtering menggunakan kesamaan antara item untuk memberikan rekomendasi kepada pengguna. Dalam proses ini, sistem memprediksirating yang mungkin akan diberikan oleh pengguna kepada item yang belum mereka nilai berdasarkan rating yang sudah mereka berikan kepada item lain yang mirip. Berikut merupakan persamaan dari prediksi:

$$P_u k = \frac{\sum_i \in_I (R_{k,i} \cdot S_{k,i})}{\sum_i \in_I |S_{k,i}|}$$

Untuk menghitung prediksi rating antara pengguna u dengan item k, untuk Rk,j dan Sk,j adalah kesamaan antara item k dan l. persamaan ini menggunakan metrik cosine similarity. Kemudian dari hasil perkalian antara item k dan l dengan rating pengguna u terhadap item k.

2.3.6 Mean Absolute Error

Untuk menilai performa suatu sistem rekomendasi dibutuhkan suatu metode dan perhitungan yang dapat mengukur tingkat kualitas prediksi yang dihasilkan oleh sistem. Mean Absolute Error (MAE) adalah salah satu metode statistika yang digunakan dalam mengevaluasi akurasi suatu sistem dengan membandingkan nilai hasil prediksi dengan nilai sesungguhnya pada data uji. Semakin rendah nilai MAE, semakin akurat prediksi yang telah dihasilkan. Berikut merupakan persamaan dari Mean Absolute Error:

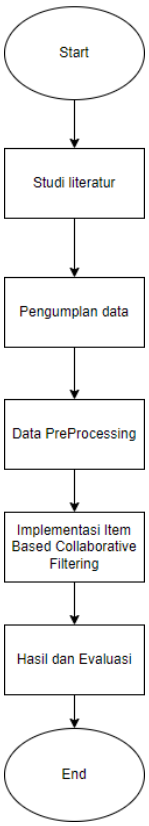
$$MAE = \frac{\sum_{i=1}^{n} |pi - qi|}{n}$$

MAE menghitung penyimpangan nilai prediksi dari nilai sesungguhnya, untuk setiap pasang nilai prediksi dan nilai sesungguhnya (pi dan qi) dituliskan dalam Persamaan.

3. METODE PENELITIAN

Pada bab ini akan dijelaskan mengenai proses- proses yang dilakukan pada penelitian penerapan algoritma Item Based Collaborative Filtering untuk dataset dalam menentukan rekomendasi produk berdasarkan rating.

Untuk proses perancangan akan dibagi beberapa tahapan yaitu studi literatur,kemudian pengumpulan data, data pre-processing, implementasi Item Based Collaborative Filtering, hasil dan evaluasi.



Gambar 3.1 Flowchart

3.1 Studi Literatur

Studi literatur digunakan untuk memperoleh informasi sebelum melakukan penelitian lebih lanjut. Hal ini juga bermanfaat untuk mengetahui metode apa yang ingin digunakan pada penelitian. Studi literatur dapat membantu untuk mengetahuimengenai teori dan konsep pada permasalahan yang diteliti seperti data mining, Item Based Collaborative Filtering, machine learning dan lain-lain. Dari studi literatur, penulis dapat mengetahui gap dari penelitian terdahulu. Untuk mengetahui berbagai teori dan konsep dapat ditemui pada jurnal penelitian nasional maupun internasional.

3.2 Pengumpulan Data

Metode pengumpulan data yang digunakan adalah metode kuantitatif. Metode kuantitatif merupakan pendekatan penelitian yang menggunakan data berupa angka atau variabel yang dapat diukur untuk mengidentifikasi pola, hubungan, dan tren. Dalam penelitian ini mengambil data sebagai penelitian yang diambil dari website Kaggle. Dataset yang digunakan pada penelitian ini adalah dataset pada toko e-commerce. Data ini memiliki beberapa fitur seperti user id, item id dan rating produk. Dataset ini berisisekitar 100.000 ribu yang berisi user id, movie id dan rating.

3.3 Data Preprocessing

Data preprocessing adalah langkah penting dalam analisis data dan pembelajaran mesin. Dua aspek penting dalam data preprocessing adalah penanganan nilai yanghilang dan duplikat. Nilai yang hilang dalam dataset bisa terjadi karena berbagai alasan,seperti kesalahan entri data, sensor yang gagal, atau data yang tidak tersedia. Adabeberapa metode untuk menangani nilai yang hilang yaitu menghapus data yang hilang.

3.4 Pembagian Data

Pembagian data adalah langkah penting dalam proses analisis data dan pembelajaran mesin. Pembagian data bertujuan untuk memastikan model yang dikembangkan dapat dievaluasi dan diuji dengan benar. Pembagian data pada penelitian ini menggunakan pembagian data training dan testing.

3.5 Implementasi Item Based Collaborative Filtering

Implementasi Item-Based Collaborative Filtering melibatkan beberapa langkah penting untuk merekomendasikan item berdasarkan kesamaan dengan item lain yang telah disukai oleh pengguna. Langkah pertama adalah membangun matriks kesamaan item, di mana setiap item dibandingkan dengan item lainnya menggunakan metrikkesamaan seperti cosine similarity atau Pearson correlation. Metrik ini mengukur seberapa mirip dua item berdasarkan pola penilaian atau interaksi pengguna terhadap kedua item tersebut. Setelah matriks kesamaan terbentuk, langkah selanjutnya adalah memilih item yang telah dinilai tinggi oleh pengguna target. Kemudian, sistem mencari item-item yang paling mirip dengan item yang sudah disukai pengguna tersebut menggunakan matriks kesamaan yang telah dibuat.

4. HASIL DAN PEMBAHASAN

4.1 Implementasi Item Based Collaborative Filtering

Implementasi Item-Based Collaborative Filtering melibatkan beberapa langkah penting untuk merekomendasikan item berdasarkan kesamaan dengan item lain yang telah

disukai oleh pengguna. Langkah pertama adalah membangun matriks kesamaan item, di mana setiap item dibandingkan dengan item lainnya menggunakan metric kesamaan seperti cosine similarity atau Pearson correlation. Metrik ini mengukur seberapa mirip dua item berdasarkan pola penilaian atau interaksi pengguna terhadap kedua item tersebut. Setelah matriks kesamaan terbentuk, langkah selanjutnya adalah memilih item yang telah dinilai tinggi oleh pengguna target. Kemudian, sistem mencari item-item yang paling mirip dengan item yang sudah disukai pengguna tersebut menggunakan matriks kesamaan yang telah dibuat.

4.2 Matriks User-Item

Matriks user-item adalah sebuah tabel dua dimensi di mana baris merepresentasikan pengguna dan kolom merepresentasikan item. Setiap sel dalam matriks berisi nilai yang menunjukkan tingkat interaksi atau preferensi pengguna terhadap item tertentu. Nilai ini bisa berupa rating. berikut ini adalah matriks user-item yang telah dibuat:

movieId	1	2	3	4	5	6	7	8
userId								
1	4.0	0.0	4.0	0.0	0.0	4.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...
606	2.5	0.0	0.0	0.0	0.0	0.0	2.5	0.0
607	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
608	2.5	2.0	2.0	0.0	0.0	0.0	0.0	0.0
609	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
610	5.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0

Matriks user-item dibuat sebagai untuk melihat interaksi antara user dengan item,kemudian mencari nilai kesamaan dengan cosine similarity. nilai tersebut dapat disimpan kedalam variabel untuk mendapatkan perhitungan dari prediksi rating.

4.3 Hasil Prediksi

Setelah mendapatkan matriks kesamaan antar item, kita dapat memprediksi rating pengguna untuk item yang belum mereka beri rating. Prediksi ini biasanya dihitung menggunakan rata-rata tertimbang dari rating item-item serupa yang telah dinilai oleh pengguna. berikut merupakan tabel hasil dari skenario uji coba prediksi darisalah satu user untuk rekomendasi film.

Tabel 4.1. Hasil Prediksi

User Id	***Movie1***	***Movie2***	***Movie3***	***Movie4***
User1	0,41	0,37	0,43	0,39
User90	0,07	0,05	0,06	0.13

5. KESIMPULAN DAN SARAN

Prediksi rating menggunakan Item-Based Collaborative Filtering telah menunjukkan hasil yang memuaskan dalam memberikan rekomendasi yang relevan kepada pengguna. Evaluasi dengan metrik MAE dan RMSE menunjukkan bahwa modelini mampu memprediksi rating dengan tingkat akurasi yang baik. Meskipun demikian, perlu adanya pendekatan tambahan untuk mengatasi tantangan seperti kelangkaan data untuk meningkatkan kinerja sistem rekomendasi.

6. DAFTAR REFERENSI

Bator, R. J., Bryan, A. D., & Schultz, P. W. (2011). Who gives a hoot?: Intercept surveys of litterers and disposers. Environment and Behavior, 43(3), 295–315. https://doi.org/10.1177/0013916509356884

Guo, J., Deng, J., Ran, X., Wang, Y., & Jin, H. (2021). An efficient and accurate recommendation strategy using degree classification criteria for item-based collaborative filtering. Expert Syst. Appl., 164, 113756. https://doi.org/10.1016/j.eswa.2020.113756

Hidayati, S. N. (2016). Pengaruh pendekatan keras dan lunak pemimpin organisasi terhadap kepuasan kerja dan potensi mogok kerja karyawan. Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship, 5(2), 57-66. http://dx.doi.org/10.30588/SOSHUMDIK.v5i2.164

Indriawan, W., Gufroni, A. I., & Informatika, J. F. T. U. S. T. (2020). Sistem rekomendasi penjualan produk pertanian menggunakan metode item based collaborative filtering. J. Siliwangi, 6(2).

Jepriana, I. W. (2018). Algoritme genetika untuk mengurangi galat prediksi metode item-based collaborative filtering. 2001, 1–7.

Jepriana, W., & Hanief, S. (2020). Analisis dan implementasi metode item-based collaborative filtering untuk sistem rekomendasi konsentrasi di STMIK STIKOM Bali. Janapati, 9(2), 171–180.

Johansson, P. (2004). Design and development of recommender dialogue systems. Institutionen för datavetenskap.

Nassar, N., Jafar, A., & Rahhal, Y. (2020). A novel deep multi-criteria collaborative filtering model for recommendation system. Knowledge-Based Syst., 187. https://doi.org/10.1016/j.knosys.2019.06.019

Rahmawati, S., Nurjanah, D., & Rismala, R. (2018). Analisis dan implementasi pendekatan hybrid untuk sistem rekomendasi pekerjaan dengan metode knowledge based dan collaborative filtering. Indones. J. Comput., 3(2), 11. https://doi.org/10.21108/indojc.2018.3.2.210

Risdwiyanto, A., & Kurniyati, Y. (2015). Strategi pemasaran perguruan tinggi swasta di Kabupaten Sleman Yogyakarta berbasis rangsangan pemasaran. Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship, 5(1), 1-23. http://dx.doi.org/10.30588/SOSHUMDIK.v5i1.142

Sallam, R. M., Hussein, M., & Mousa, H. M. (2020). An enhanced collaborative filtering-based approach for recommender systems. Int. J. Comput. Appl., 176(41), 9–15. https://doi.org/10.5120/ijca2020920531

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In Adapt. Web Lect. Notes Comput. Sci., 4321/2007, 291–324. https://doi.org/10.1007/978-3-540-72079-9_9

Shen, J., Zhou, T., & Chen, L. (2020). Collaborative filtering-based recommendation system for big data. Int. J. Comput. Sci. Eng., 21(2), 219–225. https://doi.org/10.1504/IJCSE.2020.105727

Ungkawa, U., Rosmala, D., & Aryanti, F. (2011). Pembangunan aplikasi travel recommender dengan metode case base reasoning. Jurnal Informatika, 4(1), 57–68.