Optimization of User-Based Collaborative Filtering Movie Recommendation System Using Mean-Centering and Overlap Weighting on Cosine Similarity

Arif Rohmadi [1)*)], Ery Permana Yudha [2)], Bambang Widoyono [3)]

[1)2)3)]Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret
[*)]**Correspondence author**: arifrohmadi@staff.uns.ac.id, Sukoharjo, Indonesia
DOI: https://doi.org/10.37012/jtik.v12i1.3315

Abstract

Rapid technological advancements have changed the way people enjoy entertainment. Viewers can now watch movies online without having to go to the cinema. The rapid development of online movie streaming services has increased the need for systems capable of providing content recommendations based on user preferences. This study aims to optimize a User-Based Collaborative Filtering (UBCF)-based recommendation system using the MovieLens 100K dataset containing 100,000 ratings from 943 users for 1,682 movies. The evaluation was conducted by dividing the data into 80% training data and 20% testing data. The system was optimized through two main approaches: user average normalization (mean-centering) in the KNN method to reduce bias in rating scale differences between users, and the application of overlap weighting on cosine similarity to give greater weight to user pairs with a greater number of shared item ratings. Based on experiments, a value of k = 50 was chosen as the optimal trade-off point in predicting ratings. The experimental results show that KNN with mean centering (KNNWithMeans) consistently outperforms standard KNN. The cosine + KNNWithMeans model produced an RMSE of 0.9701 and an MAE of 0.7567, lower than cosine + KNN (RMSE 1.0377; MAE 0.8226). Further overlap weighting was shown to improve prediction accuracy, with the combination of weighted cosine with α = 1 and KNNWithMeans providing the best performance with an RMSE of 0.9686 and an MAE of 0.7556.

Keywords: *Cosine Similarity, KNN, MAE, Overlap Weighting, RMSE*

Abstrak

Perkembangan teknologi yang cepat telah mengubah cara orang dalam menikmati hiburan. Penonton sekarang dapat menonton film secara online tanpa harus pergi ke bioskop. Perkembangan layanan streaming film online yang sangat pesat meningkatkan kebutuhan akan sistem yang mampu memberikan rekomendasi konten sesuai preferensi pengguna. Penelitian ini bertujuan mengoptimalkan sistem rekomendasi berbasis *User-Based Collaborative Filtering* (UBCF) menggunakan dataset MovieLens 100K yang memuat 100.000 rating dari 943 pengguna terhadap 1.682 film. Evaluasi dilakukan dengan membagi data menjadi 80% data training dan 20% data testing. Sistem dioptimalkan melalui dua pendekatan utama, yaitu normalisasi rata-rata pengguna (*mean-centering*) pada metode KNN untuk mengurangi bias perbedaan skala penilaian antar pengguna, serta penerapan *overlap weighting* pada *cosine similarity* untuk memberi bobot lebih besar pada pasangan pengguna dengan jumlah rating item bersama lebih banyak. Berdasarkan eksperimen, nilai k = 50 dipilih sebagai *trade-off point* yang optimal dalam memprediksi rating. Hasil eksperimen menunjukkan bahwa KNN dengan *mean centering* (KNNWithMeans) secara konsisten mengungguli KNN standar. Model *cosine* + KNNWithMeans menghasilkan RMSE sebesar *0.9701* dan MAE sebesar 0.7567, lebih rendah dibandingkan *cosine* + KNN (RMSE 1.0377; MAE 0.8226). Penambahan *overlap weighting* lebih lanjut terbukti mampu meningkatkan akurasi prediksi, di mana kombinasi *weighted cosine* dengan α = 1 dan KNNWithMeans memberikan performa terbaik dengan RMSE sebesar 0.9686 dan MAE 0.7556.

Kata Kunci: *Cosine Similarity*, KNN, MAE, *Overlap Weighting*, RMSE

PENDAHULUAN

Perkembangan teknologi yang cepat telah mengubah cara orang dalam menikmati hiburan. Kemunculan layanan streaming film online, mengubah preferensi penonton dan pola konsumsi. Penonton sekarang dapat menonton film secara online tanpa harus pergi ke bioskop, sehingga dapat menghemat uang tanpa harus membayar per film yang ditonton (Lai et al., 2025). Namun adanya layanan streaming juga memunculkan masalah baru berupa banyaknya konten film yang disuguhkan karena bertambah dari tahun ke tahun yang membuat pengguna kesulitan memilih film yang sesuai dengan preferensi mereka (Pandiangan et al., 2025). Oleh karena itu, perlu penerapan sistem rekomendasi yang dapat memberikan rekomendasi konten dan meningkatkan pengalaman pengguna (Shelake et al., 2025).

Sistem rekomendasi membantu pengguna menentukan pilihan dengan menyarankan konten-konten yang relevan dengan preferensi mereka (Smyth, 2025). Salah satu pendekatan dalam sistem rekomendasi adalah *Collaborative Filtering* (CF). CF memanfaatkan data interaksi pengguna dengan item untuk mengidentifikasi preferensi tersembunyi dan memberikan rekomendasi. CF cocok diterapkan pada konten seperti film maupun musik karena konten tersebut memiliki metadata yang tidak mudah untuk dideskripsikan (Nudrat et al., 2022).

Salah satu teknik dalam CF adalah *User-Based Collaborative Filtering* (UBCF). UBCF beroperasi berdasarkan prinsip bahwa pengguna memiliki minat dan perilaku yang serupa di masa lalu, cenderung memiliki preferensi yang serupa di masa depan (Permana, 2024). UBCF mencari kemiripan antar pengguna berdasarkan interaksi pengguna dengan item. Metode seperti *cosine similarity* menjadi salah satu teknik yang umum digunakan untuk mengukur kemiripan antar pengguna (Gohel & Vanjara, 2024).

Beberapa penelitian menggunakan *cosine similarity* sebagai pengukur kemiripan pada UBCF di antaranya Ma'ruf & Qoiriah (2022) melakukan perbandingan *algoritma cosine similarity* dan *euclidean distance* pada sistem rekomendasi film, hasilnya *cosine similarity* memiliki tingkat akurasi yang lebih baik. *Cosine similarity* juga digunakan pada sistem rekomendasi selain film, seperti rekomendasi produk (Suhada et al., 2023), rekomendasi konten (Gul-E-Laraib et al., 2023), dan rekomendasi restoran (Permana, 2024).

Metode *cosine similarity* tidak mempertimbangkan jumlah item yang sama-sama diberi rating oleh dua pengguna (*overlap item*) (Said et al., 2012). Oleh karena itu pada penelitian ini, kami menambahkan bobot pada perhitungan *cosine similarity*, sehingga *similarity* dari dua user dengan *overlap item* lebih banyak memiliki nilai kemiripan yang lebih tinggi. Selain itu pada proses prediksi rating, kami juga menggunakan algoritma KNN (*K-Nearest Neighbors*) yang memperhatikan rata-rata rating pengguna (*mean centering*) yang dapat mengurangi bias rating antar pengguna (Weng, 2024).

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen. Tahapan penelitian meliputi pengumpulan data, preprocessing data, penghitungan kemiripan dan rekomendasi, serta evaluasi.

Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan dataset MovieLens 100K dari GroupLens Research yang dapat diunduh melalui https://grouplens.org/datasets/movielens/. Data terdiri dari 100.000 rating oleh 943 pengguna terhadap 1.682 film. Dataset tersedia dalam format teks terpisah **.data** dan **.item**. File data memuat informasi rating pengguna yang terdiri dari user_id, movie_id, rating (1-5), dan timestamp. Adapun file item memuat informasi film yang terdiri dari movie_id, title, release_date, IMDb_URL, dan genre flag. Pada penelitian ini, data yang digunakan dari file .data hanya user_id, movie_id, dan rating seperti ditunjukkan Tabel 1.

Tabel 1. Rating Pengguna

user_id	movie_id	Rating
196	242	3
186	302	3
22	377	1
244	51	2
166	346	1

Adapun pada file .item, data yang digunakan adalah movie_id, dan title seperti ditunjukkan Tabel 2.

Tabel 2. Movie

movie_id	title
1	Toy Story (1995)
2	GoldenEye (1995)
3	Four Rooms (1995)
4	Get Shorty (1995)
5	Copycat (1995)

Preprocessing Data

Pada tahap preprocessing data dilakukan pengecekan dan pembersihan data mencakup identifikasi dan penanganan data yang tidak lengkap, data ganda, maupun data tidak valid. Proses preprocessing data diperlukan supaya data yang digunakan untuk pemodelan sudah bersih, karena kualitas data berpengaruh terhadap kualitas rekomendasi yang dihasilkan (Fitriyeh et al., 2024).

Penghitungan Kemiripan dan Rekomendasi

Data yang sudah melalui proses preprocessing selanjutnya dibagi menjadi 2, 80% data untuk proses training, dan 20% data untuk testing. Sebelum dilakukan proses penghitungan kemiripan antar pengguna, terlebih dahulu dibentuk matriks user-item. Pada matriks user-item, setiap baris mewakili pengguna, kolom mewakili film, dan nilai dalam matriks mewakili peringkat yang diberikan oleh pengguna terhadap film. Jika pengguna belum memberikan rating pada film tertentu, maka nilai rating adalah 0.

Kemiripan antar pengguna selanjutnya dihitung menggunakan *cosine similarity* dan *overlap weighted cosine similarity*. Formula *cosine similarity* ditunjukkan oleh Persamaan (i) (Al-Ghamdi et al., 2021).

$$sim(u,v) = \frac{\sum_{i \in I_{uv}} r_{ui} r_{vi}}{\sqrt{\sum_{i \in I_{uv}} r_{ui}^2} \sqrt{\sum_{i \in I_{uv}} r_{vi}^2}} \quad \text{(i)}$$

sim(u,v) menunjukkan kemiripan antara pengguna u dan v. I_{uv} merupakan himpunan film yang dirating oleh kedua pengguna. r_{ui} merupakan film yang dirating oleh pengguna u, sedangkan r_{vi} merupakan film yang dirating oleh pengguna v.

Pada penghitungan *cosine similarity* tersebut, belum mempertimbangkan banyaknya item yang dirating oleh kedua pengguna (*overlap item*). Misal user A dan user B merating 10 film yang sama dan memperoleh nilai similarity 0.8. Kemudian user A dan user C hanya

merating 5 film yang sama tetapi memperoleh nilai similarity 0.8 juga. Dari kedua kasus tersebut tidak diketahui mana pengguna yang lebih mirip preferensinya dengan pengguna A. Dengan menambahkan pembobotan (*weighting*) pada *overlap item*, maka pengguna dengan *overlap item* lebih banyak memiliki nilai *similarity* yang lebih tinggi.

Formula penambahan *overlap weighting* pada *cosine similarity* ditunjukkan persamaan (ii).

$$sim(u,v)' = sim(u,v) \times \frac{|I_{uv}|}{|I_{uv}| + \alpha} \quad \text{(ii)}$$

sim(u,v)' menunjukkan nilai kemiripan kedua pengguna yang sudah ditambahkan *weighting*. sim(u,v) merupakan nilai kemiripan pengguna menggunakan *cosine similarity*. I_{uv} merupakan jumlah film yang dirating oleh kedua pengguna. Sedangkan α merupakan pembobotan yang diberikan.

Proses selanjutnya adalah melakukan prediksi rating yang diberikan pengguna berdasarkan kedekatan dengan pengguna lainnya menggunakan formula KNN (*K-Nearest Neighbors*) *basic* seperti ditunjukkan persamaan (iii) (Li & Zhang, 2018).

$$\hat{r}_{ui} = \frac{\sum_{v \in N_i^k(u)} \text{sim}(u,v)\ r_{vi}}{\sum_{v \in N_i^k(u)} \text{sim}(u,v)} \quad \text{(iii)}$$

$\hat{r}_{ui}$ merupakan prediksi rating pengguna u terhadap film i. $N_i^k(u)$ adalah sejumlah k tetangga terdekat pengguna u yang merating film i. Sedangkan r_{vi} merupakan rating pengguna v terhadap film i.

Formula KNN *basic* belum memperhatikan kecenderungan pengguna yang sering memberikan rating tinggi terhadap semua item maupun sebaliknya. Hal tersebut tentu membuat hasil prediksi menjadi bias. Oleh karena itu pada penelitian ini juga digunakan KNNWithMeans (Weng, 2024) yaitu KNN dengan memperhatikan rata-rata rating pengguna (*mean centering*) seperti ditunjukkan pada persamaan (iv) (Al-Ghamdi et al., 2021).

$$\hat{r}_{ui} = \mu_u + \frac{\sum_{v \in N_i^k(u)} sim(u,v)\ (r_{vi} - \mu_v)}{\sum_{v \in N_i^k(u)} sim(u,v)} \quad \text{(iv)}$$

Formula KNNWithMeans mirip dengan formula KNN *basic* tetapi menambahkan rata-rata rating user (μ). Selanjutnya untuk memberikan rekomendasi ke pengguna, cukup mengambil Top-N prediksi rating dengan nilai tertinggi.

Evaluasi

Model yang telah dibuat selanjutnya dilakukan evaluasi untuk mengetahui seberapa baik kinerjanya. Evaluasi pada penelitian ini menggunakan metrik MAE *(Mean Absolute Error)*, dan RMSE *(Root Mean Squared Error)*.

MAE menghitung rata-rata selisih absolut antara rating sebenarnya dan rating hasil prediksi. Formula MAE ditunjukkan persamaan (v) (Al-Ghamdi et al., 2021).

$$MAE = \frac{1}{|\hat{R}|} \sum_{\hat{r}_{ui} \in \hat{R}} |r_{ui} - \hat{r}_{ui}| \quad \text{(v)}$$

$\hat{R}$ merupakan jumlah total rating yang diprediksi, r_{ui} merupakan nilai rating sebenarnya yang diberikan pengguna u terhadap item i, dan $\hat{r}_{ui}$ merupakan nilai rating hasil prediksi untuk pengguna u terhadap item i. Semakin kecil nilai MAE maka prediksi yang dihasilkan oleh model semakin akurat.

RMSE mirip dengan MAE, tetapi alih-alih menjumlahkan nilai absolut dari kesalahan prediksi rating, RMSE menjumlahkan kuadrat dari kesalahan tersebut menggunakan persamaan (vi) (Al-Ghamdi et al., 2021).

$$RMSE = \sqrt{\frac{1}{|\hat{R}|} \sum_{\hat{r}_{ui} \in \hat{R}} (r_{ui} - \hat{r}_{ui})^2} \quad \text{(vi)}$$

Karena menguadratkan error, maka RMSE memberi penalti lebih besar pada error besar. Nilai RMSE yang semakin kecil juga menunjukkan bahwa prediksi yang dihasilkan oleh model semakin akurat.

HASIL DAN PEMBAHASAN

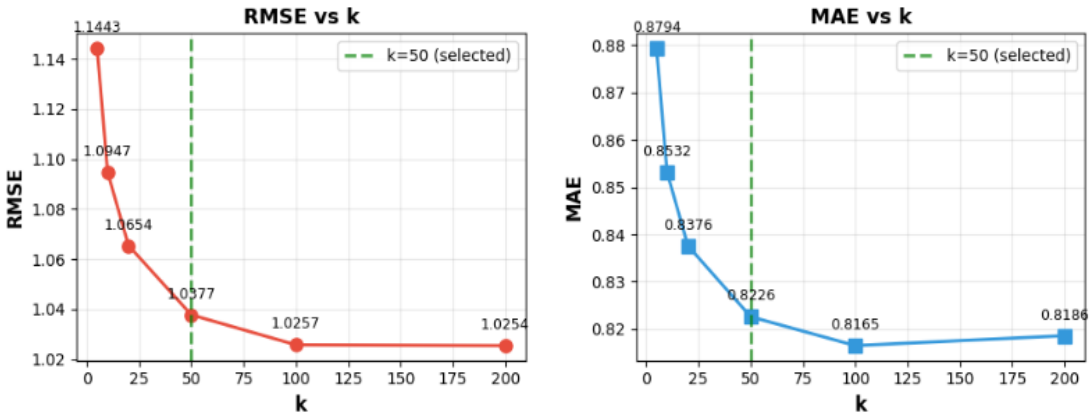
Pada penelitian ini, kami menggunakan split data (80:20) yang sudah disediakan di dataset movielens yaitu u1.base (berisi 80.000 data rating) untuk keperluan training dan u1.test (berisi 20.000 data rating) untuk keperluan testing. Dataset tersebut sudah terformat dengan baik. Ketika dilakukan preprocessing, tidak ditemukan duplikasi data, *missing value*, maupun inkonsistensi data.

Selanjutnya dibentuk matriks user-item seperti ditunjukkan Gambar 1 agar dapat dihitung kemiripan antara pengguna yang satu dengan pengguna yang lain.

movie_id	1	2	3	4	5	6	7	8	9	10	...	1673	1674	1675	1676	1677	1678	1679	1680	1681	1682
user_id																					
1	5.0	3.0	4.0	3.0	3.0	0.0	4.0	1.0	5.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	...	0.0	0.0	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	0.0	0.0	...	0.0	0.0	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	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	4.0	0.0	0.0	0.0	0.0	0.0	2.0	4.0	4.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	5.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	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	0.0	0.0	0.0	0.0	0.0	0.0
10	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	0.0	0.0	0.0	0.0

Gambar 1. Matriks User-Item

Proses penghitungan kemiripan antar pengguna menggunakan *cosine similarity* dan *weighted cosine similarity* dengan α=1,5, dan 10. Kemudian untuk prediksi rating menggunakan KNN *basic* dan KNNWithMeans. Namun, sebelumnya terlebih dahulu dilakukan eksperimen prediksi *cosine similarity* dengan KNN *basic* menggunakan variasi k yang berbeda yaitu k = 5, 10, 20, 50, 100, dan 200. Tujuannya untuk mendapatkan nilai k yang optimal untuk digunakan pada eksperimen selanjutnya. Hasil eksperimen ditunjukkan oleh Gambar 2.



Gambar 2. Perbandingan Nilai RMSE dan MAE terhadap Nilai K

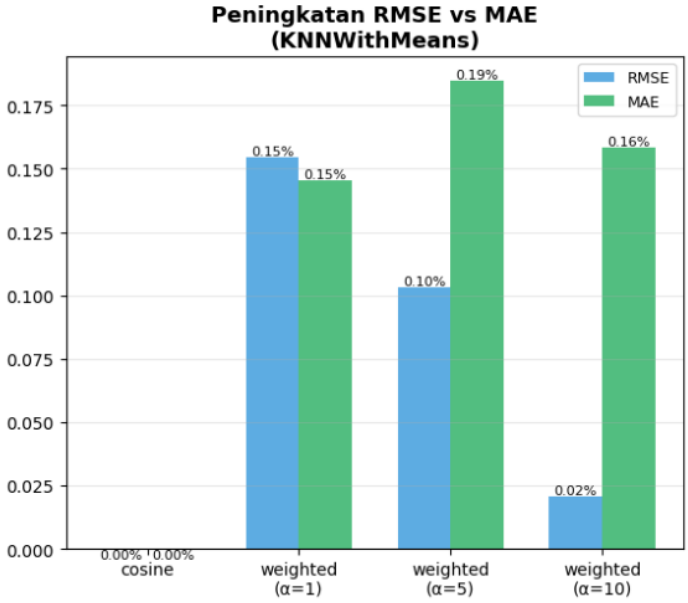
Berdasarkan Gambar 2 terlihat nilai k=50 merupakan *trade-off point* untuk nilai RMSE dan MAE, dimana untuk nilai k=100, dan k=200 peningkatan performa tidak setinggi nilai k sebelumnya.

Nilai k = 50, selanjutnya digunakan untuk menguji performa model yang lain. Hasil eksperimen ditunjukkan oleh Tabel 3.

Tabel 3. Perbandingan Performa Model

Model	RMSE	MAE
cosine + KNN	1.0377	0.8226
weighted cosine (α=1) + KNN	1.0368	0.8226
weighted cosine (α=5) + KNN	1.0351	0.8220
weighted cosine (α=10) + KNN	1.0355	0.8228
cosine + KNNWithMeans	0.9701	0.7567
weighted cosine (α=1) + KNNWithMeans	0.9686	0.7556
weighted cosine (α=5) + KNNWithMeans	0.9691	0.7553
weighted cosine (α=10) + KNNWithMeans	0.9699	0.7555

Berdasarkan Tabel 3 terlihat bahwa performa KNN dengan *mean centering* (KNNWithMeans) lebih bagus daripada KNN *basic*. RMSE untuk *cosine* + KNNWithMeans meningkat sebesar 6.52%, dan MAE meningkat sebesar 8%. Kemudian kombinasi penambahan overlap weighting terhadap KNNWithMeans memberikan peningkatan seperti ditunjukkan pada Gambar 3.



Gambar 3. Peningkatan Performa RMSE, MAE pada Kombinasi KNNWith Means

Peningkatan performa terbaik adalah ketika KNNwithMeans dikombinasikan *weighted cosine* dengan α=1. Ketika α=1 terjadi peningkatan performa RMSE dan MAE sebesar 0.15%.

Rekomendasi film yang diberikan ke pengguna, dapat diperoleh dengan mengurutkan skor film hasil prediksi model dari yang paling tinggi ke yang paling rendah. Contoh hasil rekomendasi 5 film terbaik untuk pengguna id 1 ditunjukkan Gambar 4.

```
Rekomendasi untuk user 1:
1. Schindler's List (1993)                                    (score: 16.59)
2. One Flew Over the Cuckoo's Nest (1975)                      (score: 14.2)
3. Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1963) (score: 13.67)
4. Casablanca (1942)                                           (score: 13.2)
5. Close Shave, A (1995)                                       (score: 11.67)
```

Gambar 4. Rekomendasi Film

Berdasarkan rekomendasi pada Gambar 4, terlihat bahwa model dapat menghasilkan daftar film dengan skor yang bervariasi. Hal ini menunjukkan model mampu membedakan tingkat kesesuaian antar film. Semakin tinggi nilai skor, semakin besar kemungkinan pengguna menyukai film tersebut.

KESIMPULAN DAN REKOMENDASI

Penelitian ini melakukan optimalisasi sistem rekomendasi berbasis *user-based collaborative filtering*. Berdasarkan hasil penelitian dapat disimpulkan bahwa penggunaan KNN dengan *mean centering* (KNNWithMeans) secara konsisten memberikan performa yang lebih baik dibandingkan metode KNN *basic* baik pada metrik RMSE maupun MAE. Model *cosine* + KNNWithMeans mencatat RMSE 0.9701, MAE 0.7567 jauh lebih rendah dibandingkan *cosine* + KNN dengan RMSE 1.0377, dan MAE 0.8226.

Selain itu penambahan *overlap weighting* pada *cosine similarity* terbukti meningkatkan akurasi prediksi. Nilai α=1 dikombinasikan dengan KNNWithMeans menghasilkan performa terbaik dengan RMSE terendah sebesar 0.9686.

Sejalan dengan SDG ke 9 (industri, inovasi, dan infrastruktur), penelitian ini berkontribusi pada pengembangan infrastruktur digital yang lebih cerdas untuk layanan rekomendasi. Kedepan, kinerja model dapat ditingkatkan melalui uji signifikansi berbasis *k-fold*, eksplorasi pemodelan *hibrida*, maupun eksperimen pada data lain untuk meningkatkan keandalan dan generalisasi pada skenario data yang bervariasi.

UCAPAN TERIMA KASIH

Penelitian ini didanai oleh RKAT Universitas Sebelas Maret Tahun Anggaran 2026 melalui skema Penelitian PENGUATAN KAPASITAS GRUP RISET (PKGR-UNS) C dengan Nomor Perjanjian Penugasan Penelitian : 462/UN27.22/PT.01.03/2026.

REFERENSI

Al-Ghamdi, M., Elazhary, H., & Mojahed, A. (2021). Evaluation of Collaborative Filtering for Recommender Systems. International Journal of Advanced Computer Science and Applications, 12(3). https://doi.org/10.14569/IJACSA.2021.0120367

Fitriyeh, F., Haqqi, N., Choiriyah, L. M., & Ifada, N. (2024). Dampak pembobotan pada metode hybrid user-based dan item-based untuk sistem rekomendasi film. *Jurnal CoSciTech (Computer Science and Information Technology)*, *5*(3), 516–525.

Gohel, D. D. P., & Vanjara, D. P. A. (2024). "Analyzing User-Based and Item-Based Recommender Systems: A Comparative Examination." *Educational Administration: Theory and Practice*, *30*(6(S)), 91–98. https://doi.org/10.53555/kuey.v30i6(S).5331

Gul-E-Laraib, Zaman, S. K. uz, Maqsood, T., Rehman, F., Mustafa, S., Khan, M. A., Gohar, N., Algarni, A. D., & Elmannai, H. (2023). Content Caching in Mobile Edge Computing Based on User Location and Preferences Using Cosine Similarity and Collaborative Filtering. *Electronics*, *12*(2), 284. https://doi.org/10.3390/electronics12020284

Lai, C.-Y., Hashim, N., & Krisnawati, L. D. (2025). Hybrid-Based Movie Recommender System: Techniques, Case Studies, Evaluation Metrics, and Future Trends. *Journal of Informatics and Web Engineering*, *4*(3), 90–110. https://doi.org/10.33093/jiwe.2025.4.3.5

Li, G., & Zhang, J. (2018). Music personalized recommendation system based on improved KNN algorithm. *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 777–781. https://doi.org/10.1109/IAEAC.2018.8577483

Ma'ruf, M. A., & Qoiriah, A. (2022). Perbandingan Algoritma Cosine Similarity dan Euclidean Distance pada Sistem Rekomendasi Film dengan Metode Item-Based Collaborative Filtering. *Journal of Informatics and Computer Science (JINACS)*, 160–168. https://doi.org/10.26740/jinacs.v4n02.p160-168

Nudrat, S., Khan, H. U., Iqbal, S., Talha, M. M., Alarfaj, F. K., & Almusallam, N. (2022). Users' Rating Predictions Using Collaborating Filtering Based on Users and Items Similarity Measures. *Computational Intelligence and Neuroscience*, *2022*(1), 2347641. https://doi.org/10.1155/2022/2347641

Pandiangan, A. Z., Huliman, & Pribadi, O. (2025). Design of a Movie Recommendation System Application Using a User-Based Collaborative Filtering Algorithm on Android. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, *5*(1), 1196–1200. https://doi.org/10.59934/jaiea.v5i1.1583

Permana, K. E. (2024). *Comparison of User Based and Item Based Collaborative Filtering in Restaurant Recommendation System | IIETA*. IIETA. https://doi.org/10.18280/mmep.110723

Said, A., Jain, B. J., & Albayrak, S. (2012). Analyzing weighting schemes in collaborative filtering: Cold start, post cold start and power users. *Proceedings of the 27th Annual ACM Symposium on Applied Computing, SAC '12*, 2035–2040. https://doi.org/10.1145/2245276.2232114

Shelake, V., Fernandes, S., & Shrungare, S. (2025). AI-Driven Personalized Movie Recommendations: A Content and Sentiment-Aware Model for Streaming and Digital Entrepreneurship. *Aptisi Transactions on Technopreneurship (ATT)*, *7*(2), 306–317. https://doi.org/10.34306/att.v7i2.550

Smyth, B. (2025). People Who Liked This Also Liked ... A Publication Analysis of Three Decades of Recommender Systems Research. *ACM Trans. Recomm. Syst.*, *4*(2), 28:1-28:41. https://doi.org/10.1145/3742442

Suhada, S., Bahri, S., Nugraha, S. B., Hidayatulloh, T., & Wintana, D. (2023). Sistem Rekomendasi Produk Menggunakan Metode User-Based Collaborative Filtering Pada Digital Marketing. *J-INTECH*, *11*(1), 158–166. https://doi.org/10.32664/j-intech.v11i1.866

Weng, J. (2024). Comparison of machine learning-based book comments sentiment analysis for constructing recommendation system. *Applied and Computational Engineering*, *47*, 120–124. https://doi.org/10.54254/2755-2721/47/20241245