



Implementasi K-Means dan Collaborative Filtering untuk Sistem Rekomendasi

Melany Mustika Dewi✉, Lilis Dwi Farida, dan Akhmad Dahlan

Jurusan Manajemen Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, Indonesia

Info Artikel

Sejarah Artikel:

Diterima: 29 Agustus 2023

Direvisi: 24 Juli 2024

Disetujui: 27 Juli 2024

Keywords:

Collaborative Filtering,
Item Based Collaborative
Filtering, Sistem
Rekomendasi, K-Means

Abstrak

Pada penelitian ini akan mengadopsi K-Means untuk mengelompokkan film yang serupa sebelum dilakukan rekomendasi menggunakan *item based collaborative filtering* dengan tujuan mereduksi data, mendapatkan data kelompok film yang sesuai dengan target pengguna dan mempercepat proses eksekusi. Penelitian ini dimulai dengan melakukan pengumpulan data, pengolahan data, proses klastering, proses rekomendasi, proses prediksi rating dan akurasi. Dari hasil penelitian dapat disimpulkan bahwa penggunaan metode *item based collaborative filtering* pada dataset MovieLens 100k dapat memberikan rekomendasi film kepada pengguna. Namun akurasi MAE dari prediksi rating memiliki error sedang yaitu 0.54 dan akurasi menggunakan metode RMSE memiliki nilai error yang tinggi yaitu 0.71. Waktu eksekusi untuk proses rekomendasi juga memakan waktu lama yaitu 14 menit 12 detik.

Abstract

In this study, K-Means will be adopted to group similar films before making recommendations using item-based collaborative filtering with the aim of reducing data, obtaining data for groups of films that are in accordance with the target user and speeding up the execution process. This research begins with data collection, data processing, clustering process, recommendation process, rating prediction process and accuracy. From the results of the study it can be concluded that the use of the item-based collaborative filtering method on the MovieLens 100k dataset can provide film recommendations to users. However, the MAE accuracy of the rating prediction has a moderate error of 0.54 and the accuracy using the RMSE method has a high error value of 0.71. The execution time for the recommendation process also takes a long time, namely 14 minutes 12 seconds.

PENDAHULUAN

Pada era saat ini film telah menjadi salah satu bentuk hiburan favorit utama di kalangan masyarakat. Berdasarkan data dari Badan Pusat Statistik, Jumlah film yang ditayangkan dibioskop Indonesia pada tahun 2018 mencapai 100.000 film (Statistik, 2018). Dengan kondisi seperti ini mengakibatkan situasi di mana para penggemar film menghadapi kesulitan dalam mencari informasi yang cocok dengan preferensi mereka. Film-film yang tidak pernah terlintas dalam pikiran mereka atau tidak sesuai dengan kebiasaan yang mereka tonton terkadang memiliki daya tarik dan sesuai dengan selera mereka. Salah satu solusi dari tantangan ini adalah sistem rekomendasi sebagai teknologi untuk memberikan rekomendasi dalam suatu keputusan yang menggunakan opini dan penilaian orang lain terhadap sebuah film. (Muarif & Winarno, 2022)

Sistem rekomendasi adalah algoritma yang mampu memberikan rekomendasi mengenai suatu item sesuai dengan personalisasi pengguna dengan harapan rekomendasi tersebut dapat menghemat waktu pencarian item dan membantu pengguna dalam mengambil keputusan (Dharma et al., 2021; Jaja et al., 2020; Yoshua et al., 2021). Pengertian lainnya, Sistem rekomendasi adalah perangkat lunak yang mempelajari minat dan preferensi pengguna terhadap suatu item dan memberikan rekomendasi yang sesuai dengan kebutuhan pengguna tersebut (Islamiyah, Subekti, Dwi Andini, et al., 2019). Sistem rekomendasi diklasifikasikan menjadi tiga jenis yaitu *Content Based Filtering*, *Collaborative Filtering* dan *Hybrid Filtering* (Salloum & Rajamanthri, 2021). *Content Based Filtering* adalah jenis sistem rekomendasi yang memberikan rekomendasi menggunakan deskripsi dari suatu item (Widiyaningtyas et al., 2021). *Collaborative Filtering* memberikan rekomendasi dengan menghitung kemiripan berdasarkan preferensi dari pengguna (AL-Bakri & Hashim, 2019). *Collaborative Filtering* mempunyai dua tipe yaitu *User based collaborative filtering* merekomendasikan suatu item yang disukai pengguna lain yang serupa kemiripan preferensinya dan *Item based collaborative filtering* memberikan rekomendasi berdasarkan produk yang memiliki kemiripan dengan produk yang dinilai atau disukai oleh pengguna tersebut (Sharma & Yadav, 2020). *Hybrid Filtering* adalah sistem rekomendasi yang menggunakan algoritma gabungan dari *Content Based Filtering* dan *Collaborative Filtering* (Singh et al., 2020).

Peneliti akan menggunakan Metode *Item Based Collaborative filtering* yang banyak diadopsi dalam sistem rekomendasi industri karena memiliki

model yang mudah digunakan untuk personalisasi online, mudah diterapkan dalam suatu sistem dan mampu mengeksplorasi asosiasi yang berasal dari riwayat pengguna target disebut asosiasi implisit sehingga meningkatkan ketepatan hasil rekomendasi (Islamiyah, Subekti, & Andini, 2019).

Penelitian lain yang telah dilakukan adalah menggunakan *item based collaborative filtering* untuk memberikan rekomendasi aksesoris *smartphone* dengan mencari kesamaan item dengan item lain menggunakan *pearson similiarity* (Prasetyo et al., 2019). Evaluasi dari hasil rekomendasi dilakukan dengan menggunakan algoritma *Mean Absolute Error* (MAE). Data yang digunakan memiliki rating yang minim. Hasil dari penelitian tersebut adalah sistem dapat memberikan rekomendasi *smartphone* dengan nilai error kecil yaitu 0,57, akurasi yang baik didapat dari data konsumen yang memiliki kesamaan karakteristik. Kelemahan dari penelitian ini adalah waktu untuk proses eksekusi program cukup lama yaitu 6,4 detik.

Peneliti lain menggunakan *item based collaborative filtering* untuk merekomendasikan konsentrasi di STMIK STIKOM BALI menggunakan data nilai mahasiswa Program Studi Sistem Komputer dan Sistem Informasi (Jepriana & Hanief, n.d.). Hasil dari pengujian mendapatkan nilai *precision* sebesar 0,98 yaitu terbukti dengan data indeks prestasi mata kuliah konsentrasi lebih atau sama dengan 2,75 akan direkomendasikan sehingga dapat disimpulkan bahwa sistem cukup baik dalam memberikan rekomendasi. Kekurang dari penelitian ini adalah perlu meningkatkan relevansi dari hasil rekomendasi.

Peneliti lain telah berhasil mengatasi skalabilitas data pada sistem rekomendasi *user-based collaborative filtering* menggunakan algoritma K-Means pada matriks rating film untuk menemukan pengguna serupa akan dikelompokkan dalam cluster yang sama (AL-Bakri & Hashim, 2019). Penelitian ini menggunakan dataset MovieLens 100k yang memiliki 100.000 rating dari 943 pengguna di 1682 film dengan mengambil kolom *the userid*, *movieid* dan *rating*. Hasil yang diperoleh dengan menerapkan *K-Means Clustering* pada data MovieLens mengurangi ketersebaran dan masalah skalabilitas pada sistem rekomendasi. Penggunaan *Pearson Correlation* mencapai hasil yang relatif baik untuk menemukan item terdekat dengan target pengguna.

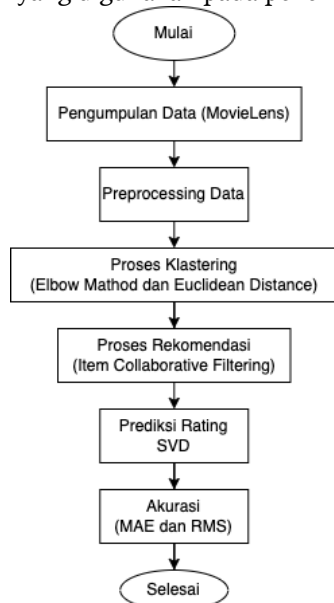
K-Means merupakan salah satu algoritma untuk mengelompokkan data yang memiliki karakteristik yang sama (Apriyani et al., 2023).

Kelebihan dari algoritma ini adalah efisien dan menghasilkan hasil yang optimal namun memiliki kekurangan yaitu hasil bergantung dari jumlah kelompok yang ditentukan (Nugraha et al., 2021). Untuk menentukan jumlah kelompok yang sesuai pada algoritma K-Means dapat menggunakan metode elbow (Wahyudi et al., 2021).

Pada penelitian ini akan mengadopsi K-Means untuk mengelompokkan film yang serupa sebelum dilakukan rekomendasi menggunakan *item based collaborative filtering*. Tujuan dari menggunakan K-Means adalah mereduksi data, mendapatkan data kelompok film yang sesuai dengan target pengguna dan mempercepat proses eksekusi. Hasil dari penelitian ini adalah rekomendasi film yang sesuai dengan target pengguna menggunakan K-Means dan *item based collaborative filtering* serta hasil pengujian dari kolaborasi algoritma tersebut.

METODE PENELITIAN

Berikut ini pada Gambar 1 adalah metode penelitian yang digunakan pada penelitian ini.



Gambar 1. Metodologi Penelitian

A. Pengumpulan Data

Tahapan pertama dari metode penelitian adalah melakukan pengumpulan data. Penelitian ini menggunakan dataset MovieLens 100k yang didapat diunduh dari melalui <https://grouplens.org/datasets/movielens/>. Dataset MovieLens memiliki 100.000 rating, 3600 tag pada 9000 film dari 600 pengguna. Terdapat 4 file csv pada dataset MovieLens yaitu *movies*, *ratings*, *tags* dan *links*. Dataset ini diperbaharui pada bulan Agustus tahun 2018.

B. Preprocessing Data

Tahapan kedua adalah melakukan *preprocessing* data dengan menentukan data yang akan digunakan pada penelitian dan menyesuaikan data dengan algoritma yang digunakan. Dataset MovieLens memiliki 4 file yaitu *movies*, *ratings*, *tags* dan *links*. Peneliti menggunakan dua file yaitu *movies* dan *rating*. File *movies* dengan kolom *movieid* yang berisi id dari film, kolom *title* berisi judul film, dan kolom *genres* berisi kategori film. File *Movies* memiliki data sebanyak 193.609 baris. Pada Tabel 1 ditunjukkan data dari file *Movies*.

Tabel 1. Data Movies

movie id	title	genres
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Men (1995)	Comedy Romance
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy
6	Heat (1995)	Action Crime Thriller
7	Sabrina (1995)	Comedy Romance
8	Tom and Huck (1995)	Adventure Children
193.609	GoldenEye (1995)	Action Adventure Thriller

File *Ratings* memiliki kolom *userid* yang berisi id dari pengguna, kolom *movieid* berisi id dari film, kolom *rating* berisi nilai rating yang diberikan oleh pengguna dan kolom *timestamp* berisi kode waktu dan tanggal saat pemberian rating oleh pengguna. File *Ratings* memiliki data sebanyak 100.836 rating dari 610 pengguna. Pada Tabel 2 ditunjukkan data dari file *Ratings*.

Tabel 2. Data Ratings

userid	movieid	rating	timestamp
1	1	4	964982703
1	3	4	964981247
1	6	4	964982224
1	47	5	964983815
1	50	5	964982931
1	70	3	964982400
1	101	5	964980868
1	110	4	964982176
610	157	5	964984100

Akan dilakukan penggabungan File *Movies* dan *Ratings* dan diambil kolom yang diperlukan dalam penelitian yaitu *userid*, *movieid*, *title*, *genres* dan *rating* seperti yang ditunjukkan pada Tabel 3.

Tabel 3. Data Gabungan

user id	movie id	title	genres	rating
1	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	4
1	3	Grumpier Old Men (1995)	Comedy Romance	4
1	6	Heat (1995)	Action Crime Thriller	4
1	47	Seven (a.k.a. Se7en) (1995)	Mystery Thriller	5
1	50	Usual Suspects, The (1995)	Crime Mystery Thriller	5
1	70	From Dusk Till Dawn (1996)	Action Comedy Horror Thriller	3
1	110	Braveheart (1995)	Action Drama War	4
1	157	Canadian Bacon (1995)	Comedy War	5

Tahapan selanjutnya adalah melakukan *encoding* kolom *genres* yang bertipe data *string* menjadi kode atau integer agar dapat dilakukan proses klastering.

C. Proses Klastering (K-Means)

Proses klastering atau pengelompokan film berdasarkan kolom *movieid*, *rating* dan *genres* yang sudah dilakukan *encoding* menggunakan algoritma K-Means. Berikut langkah penerapan algoritma K-Means sebagai berikut (Mirantika, 2021):

1. Pilihlah jumlah *k* yang akan digunakan sebagai jumlah klaster atau kelompok dengan menggunakan *elbow method*.
2. Pilih centroid atau nilai tengah dari masing – masing klaster secara acak.
3. Hitung jarak dari tiap data terhadap centroid menggunakan algoritma *Euclidean Distance*. Berikut pada persamaan 1 adalah rumus dari algoritma K-Means.

$$D(x_i, c_j) = \sqrt{\sum_{i=1}^N \sqrt{(x_i - c_j)^2}} \quad (1)$$

Keterangan :

$D(x_i, c_j)$ = Jarak dari *i* ke pusat cluster *j*
 x_i = Data ke *i* pada atribut data ke *k*
 c_j = Titik pusat ke *j* pada atribut ke *k*

4. Masukkan data ke dalam klaster sesuai jarak terdekatnya yang dihasilkan dari langkah c.
5. Ulangi hingga data klaster tidak berubah.

D. Proses Rekomendasi

Rekomendasi dilakukan dengan metode item based collaborative filtering dengan tahapan sebagai berikut :

1. Pembuatan matriks kemiripan dari hasil rating + hasil encoding genres dengan nama hasil menggunakan metode *cosine similarity*.
2. Pencarian film yang memiliki hasil kemiripan tinggi.

E. Prediksi Rating

Bedasarkan data *movieid* dan *userid* akan dilakukan prediksi rating untuk mengatasi film yang belum memiliki rating menggunakan algoritma *Singular Value Decomposition* (SVD) akan mengatasi kelemahan tersebut (Al Amin et al., 2020). Berikut pada persamaan 2 adalah rumus SVD.

$$R = M(\sum U^t) \quad (2)$$

Keterangan :

R = matriks rating

M = *eigenvector* matriks RR^T

U = *eigenvector* matriks R^TR

F. Akurasi

Evaluasi dilakukan dengan menggunakan metode *Mean Absolute Error* (MAE) dan *Root Mean Square* (RMS). Pada persamaan 3 merupakan rumus RMS.

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (3)$$

Keterangan :

MAE = *mean absolute error*

y_i = prediksi

x_i = nilai asli

n = total data

Berikut pada persamaan 4 disajikan rumus RMS.

$$RMS = \sqrt{\frac{1}{n} \sum x_i^2} \quad (4)$$

Keterangan :

RMS = *root mean square*

n = jumlah data

x_i = nilai

HASIL DAN PEMBAHASAN

A. Preprocessing Data

Hasil dari pengolahan data dalam terbentuknya data yang akan digunakan menggunakan aplikasi Google Colaboratory

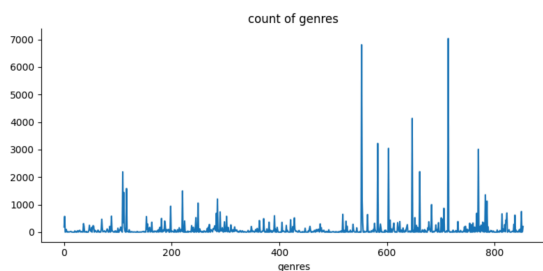
dengan RAM 12 GB dan Bahasa pemrograman Python 3. Berikut pada Gambar 2 adalah keseluruhan data yang digunakan dengan kolom *movieid*, *title*, *userid*, *rating* dan *genres*.

	movieId	title	userId	rating	genres
0	1	Toy Story (1995)	3	4.0	302
1	1	Toy Story (1995)	6	5.0	302
2	1	Toy Story (1995)	8	4.0	302
3	1	Toy Story (1995)	10	4.0	302
4	1	Toy Story (1995)	11	4.5	302
...
99995	119141	The Interview (2014)	631	3.0	153
99996	125916	Fifty Shades of Grey (2015)	31	0.5	714
99997	125916	Fifty Shades of Grey (2015)	692	0.5	714
99998	128488	Wild Card (2015)	284	3.0	683
99999	128594	Boy Meets Girl (2015)	692	0.5	603

100000 rows x 5 columns

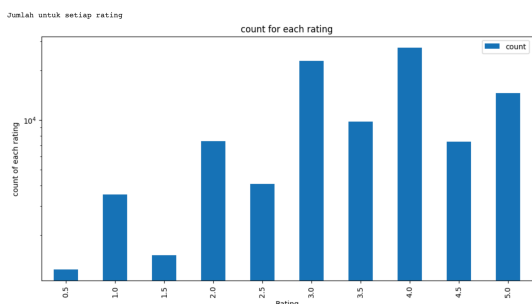
Gambar 2. Data rekomendasi

Pada gambar 2 menampilkan sebaran dari data *genres* dalam bentuk diagram batang.



Gambar 3. Sebaran data *genres*

Pada gambar 4 menampilkan diagram batang dari sebaran data *rating*.



Gambar 4. Sebaran data *rating*

Dari hasil dari pengecekan data tidak terdapat kolom yang tidak terisi atau kosong (*null*) seperti yang di paparkan pada Gambar 5.

```
data.isnull().sum()

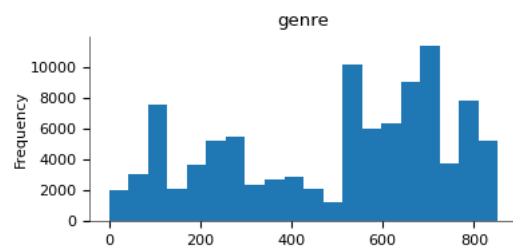
movieId    0
title      0
userId     0
rating     0
genres     0
dtype: int64
```

Gambar 5. Hasil pengecekan data *null*

Pengecekan ketidaklengkapan data atau sering disebut *sparsity data* dari pengguna yang belum melakukan rating atau penilaian film mencapai 98,3%, yang termasuk dalam kategori tinggi.

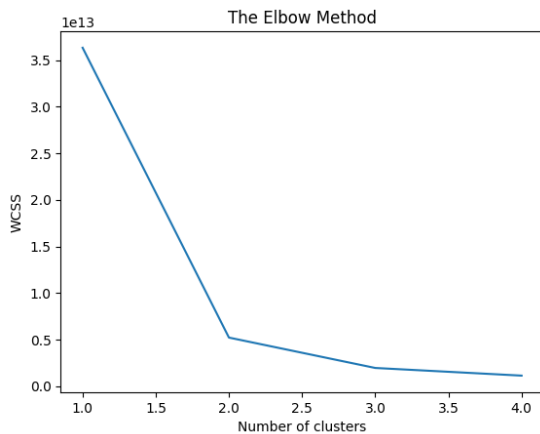
B. Proses Klustering (K-Means)

Langkah pertama dari proses ini adalah melakukan encoding dari kolom *genres* yang memiliki tipe data *string* diubah menjadi kode yang memiliki tipe data *integer*. Berikut pada Gambar 6 adalah sebaran data dari hasil *encoding* kolom *genres*.



Gambar 6. Sebaran data *encoding* kolom *genres*

Langkah selanjutnya adalah menggunakan data dengan kolom *movieid*, *userid*, *rating* dan *genres* yang sudah dikode untuk dilakukan proses klustering dengan algoritma K-Means. Penentuan jumlah kluster menggunakan metode *elbow* dengan hasil 2 kluster seperti pada Gambar 7.

Gambar 7. Hasil jumlah kluster dengan metode *elbow*

Berikut pada Tabel 4 menyajikan hasil dari Kluster 1 dan Tabel 5 menyajikan hasil dari Kluster 2 dari proses klustering yang telah dilakukan.

Tabel 4. Hasil Kluster 1

movieid	userid	rating	genres
1	3	4.0	302
1	6	5.0	302
1	8	4.0	302
1	10	4.0	302
1	11	4.5	302
...
32058	572	3.0	714
32078	480	3.0	242
32116	91	2.5	553
32116	132	2.5	553

Tabel 5. Hasil Kluster 2

movieid	userid	rating	genres
32139	91	3.0	714
32141	359	3.5	714
32160	208	4.5	553
32170	359	4.5	661
32179	394	4.0	683
...
119141	631	3.0	153
125916	31	0.5	714
125916	692	0.5	714
128488	284	3.0	683

C. Proses Rekomendasi

Berikut pada Gambar 8 hasil matrik kemiripan menggunakan *cosine similarity* berdasarkan nilai hitung ($\text{rating} + \text{encoding genres}$).

	userid	1	2	3	4	5	6	7	8	9	10	11
title												
2001: A Space Odyssey (1968)		55.5	57.0	57.0	0	0	0	55.0	0	0	0	0
Affair to Remember, An (1957)		0	0	0	0	0	0	91.0	0	0	0	0
African Queen, The (1951)		0	0	0	0	0	0	0	0	0	51.0	0
Aladdin (1992)		0	0	0	0	43.0	0	0	41.0	0	0	0
All Dogs Go to Heaven 2 (1996)		0	0	0	0	43.0	0	0	0	0	0	0
...	
Truth About Cats & Dogs, The (1996)		0	0	0	0	76.0	77.0	76.0	0	0	0	0
Twister (1996)		0	0	0	0	15.0	12.0	13.0	0	0	0	0
Visitors, The (Visiteurs, Les) (1993)		0	0	74.0	0	0	0	0	0	0	0	0
Wallace & Gromit: The Best of Aardman Animation (1996)		0	0	0	0	46.0	0	0	0	0	0	0
Wizard of Oz, The (1939)		47.5	0	48.0	0	0	0	0	0	0	0	0

Gambar 8. Matriks *similarity*

Selanjutnya adalah melakukan percobaan nilai kemiripan film menggunakan nilai dari matriks, Hasil kemiripan antara film *Snow White and the Seven Dwarfs* (1937) dengan *Secret Garden, The* (1993) adalah 0.32. Dapat dikatakan bahwa kedua film tersebut kurang mirip atau tidak memiliki preferensi rating dan *genres* yang sama.

Hasil rekomendasi film *Snow White and Seven Dwarfs* (1937) dari 100.000 data seperti pada Gambar 9 dengan waktu eksekusi yang lama yaitu 14 menit 12 detik.

0	0.568632	Pinocchio (1940)
1	0.509517	Cinderella (1950)
2	0.502724	Lion King, The (1994)
3	0.475862	Beauty and the Beast (1991)
4	0.453226	Wizard of Oz, The (1939)
5	0.451193	Aladdin (1992)
6	0.448774	Home Alone (1990)
7	0.439587	Fantasia (1940)
8	0.431063	Mary Poppins (1964)
9	0.421172	Jungle Book, The (1967)

Gambar 9. Hasil rekomendasi sistem

D. Prediksi Rating

Sebelumnya melakukan *install library surprise* pada Google Colaboratory. Dengan menggunakan algoritma SVD dan dilakukan 5 kali *cross validation* didapatkan hasil prediksi rating seperti pada Gambar 10.

movieid	userid	rating	genres_x	hitung	title	PrediksiRating
0	1	3	4.0	302	Toy Story (1995)	4.505310
1	1	6	5.0	302	Toy Story (1995)	4.284723
2	1	8	4.0	302	Toy Story (1995)	4.139422
3	1	10	4.0	302	Toy Story (1995)	3.750669
4	1	11	4.5	302	Toy Story (1995)	4.855508
...
90019	32058	572	3.0	714	Class Action (1991)	3.402198
90020	32078	480	3.0	242	Godzilla vs. Mechagodzilla II (Gojira VS Mekag...	3.099075
90021	32116	91	2.5	553	Oh, God! You Devil (1984)	2.974559
90022	32116	132	2.5	553	Oh, God! You Devil (1984)	3.035099
90023	32116	664	3.5	553	Oh, God! You Devil (1984)	3.839595

Gambar 10. Hasil prediksi rating

E. Evaluasi

Evaluasi dilakukan dengan dua metode, MAE dan RMSE dengan data menggunakan rating asli dan prediksi rating. Pada Gambar 11 terdapat kode program untuk mencari nilai MAE dan menghasilkan nilai error 0.54.

```
1 from sklearn.metrics import mean_absolute_error
  y_true = cluster.rating
  y_pred = cluster.PrediksiRating
  mean_absolute_error(y_true, y_pred)

0.5498206994264634
```

Gambar 11. Hasil akurasi MAE

Pada Gambar 12 terdapat kode program dan nilai error menggunakan RMSE. Hasil dari nilai error dapat dikatakan tinggi yaitu 0.71 hampir mendekati 1.

```
1 from sklearn.metrics import mean_squared_error
  from math import sqrt
  y_true = cluster.rating
  y_pred = cluster.PrediksiRating
  rms = sqrt(mean_squared_error(y_true, y_pred))
  rms

0.7102948924734104
```

Gambar 12. Hasil akurasi RMSE

SIMPULAN

Dari hasil penelitian dapat disimpulkan bahwa penggunaan metode item based collaborative filtering pada dataset MovieLens 100k dapat memberikan rekomendasi kepada pengguna. Namun akurasi MAE dari rating yang diberikan oleh sistem memiliki error cukup sedang yaitu 0.54, semakin dekat dengan angka 1 maka semakin tinggi errornya. Nilai menggunakan metode RMSE dapat dikatakan tinggi yaitu 0.71. Waktu eksekusi untuk rekomendasi termasuk kategori lama yaitu 14 menit 12 detik. Saran dari peneliti untuk penelitian selanjutnya dapat menggunakan aplikasi lain dengan RAM yang lebih tinggi atau menggunakan algoritma lainnya untuk mempersingkat waktu eksekusi dan meningkatkan akurasi dari prediksi rating.

DAFTAR PUSTAKA

- Al Amin, A., Sunyoto, A., & Al Fatta, H. (2020). Mereduksi Error Prediksi Pada Sistem Rekomendasi Menggunakan Pendekatan Collaborative Filtering Berbasis Model Matrix Factorization. *EXPLORE*, 10(1).
- AL-Bakri, N. F., & Hashim, S. H. (2019). Collaborative Filtering Recommendation Model Based on k-means Clustering. *Al-Nahrain Journal of Science*, 22(1), 74–79. <https://doi.org/10.22401/ANJS.22.1.10>
- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20–33. <https://doi.org/10.56211/helloworld.v2i1.230>
- Dharma, A. S., Basadena, R. B., Hutasoit, A., & Pangaribuan, R. R. (2021). *Arie Satia Dharma: Sistem Rekomendasi Menggunakan Item-based ... Sistem Rekomendasi Menggunakan Item-based Collaborative Filtering pada Konten Artikel Berita Sejarah penerimaan*.
- Islamiyah, M., Subekti, P., & Andini, T. D. (2019). Utilization of CollaboratPemanfaatan Metode Based Collaborative Filtering Untuk Rekomendasi Wisata Di Kabupaten Malangive Filtering Method for Tourism Recommendations in Malang Regency. *Jurnal Ilmiah Teknologi Informasi Asia*, 13(2), 143. <https://doi.org/10.32815/jitika.v13i2.70>
- Islamiyah, M., Subekti, P., Dwi Andini, T., & Asia Malang, S. (2019). Pemanfaatan Metode Item Based Collaborative Filtering Untuk Rekomendasi Wisata Di Kabupaten Malang. *Jurnal Ilmiah Teknologi Informasi Asia*, 13(2).
- Jaja, V. L., Susanto, B., & Sasongko, L. R. (2020). Penerapan Metode Item-Based Collaborative Filtering Untuk Sistem Rekomendasi Data MovieLens. *D'CARTESIAN*, 9(2), 78. <https://doi.org/10.35799/dc.9.2.2020.28274>
- Jepriana, W., & Hanief, S. (n.d.). Analisis Dan Implementasi Metode Item-Based Collaborative Filtering Untuk Sistem Rekomendasi Konsentrasi Di Stmik Stikom Bali (Vol. 9, Issue 2).
- Mirantika, N. (2021). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 di Provinsi Jawa Barat. *NUANSA INFORMATIKA*, 15(2), 92–98. <https://doi.org/10.25134/nuansa.v15i2.4321>
- Muarif, A. S., & Winarno, E. (2022). Sistem Rekomendasi Tempat Parkir di Kota Lama Semarang Menggunakan

- Collaborative Filtering. *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 906. <https://doi.org/10.33087/jiubj.v22i2.2066>
- Nugraha, D., Purboyo, T. W., & Nugrahaeni, R. A. (2021). *Sistem Rekomendasi Film Menggunakan Metode User Based (Movie Recommendation System Using User Based Collaborative Filtering Method)*. 8(5), 6765–6775.
- Prasetyo, B., Haryanto, H., Astuti, S., Astuti, E. Z., & Rahayu, Y. (2019). Implementasi Metode Item-Based Collaborative Filtering dalam Pemberian Rekomendasi Calon Pembeli Aksesoris Smartphone. *Eksplora Informatika*, 9(1), 17–27. <https://doi.org/10.30864/eksplora.v9i1.244>
- Salloum, S., & Rajamanthri, D. (2021). Implementation and evaluation of movie recommender systems using collaborative filtering. *Journal of Advances in Information Technology*, 12(3), 189–196. <https://doi.org/10.12720/jait.12.3.189-196>
- Sharma, P., & Yadav, L. (2020). Movie Recommendation System Using Item Based Collaborative Filtering. *International Journal of Innovative Research in Computer Science & Technology*, 8(4). <https://doi.org/10.21276/ijircst.2020.8.4.2>
- Singh, R. H., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. (2020). Movie Recommendation System using Cosine Similarity and KNN. *International Journal of Engineering and Advanced Technology*. <https://doi.org/10.35940/ijeat.E9666.069520>
- Statistik, B. P. (2018). *Persentase Judul Film yang Ditayangkan oleh Perusahaan Bioskop Menurut Genre (Persen), 2014-2018*. Badan Pusat Statistik.
- Wahyudi, I., Sulthan, M. B., & Suhartini, L. (2021). Analisa Penentuan Cluster Terbaik Pada Metode K-Means Menggunakan Elbow Terhadap Sentra Industri Produksi Di Pamekasan. *Jurnal Aplikasi Teknologi Informasi Dan Manajemen (JATIM)*, 2(2), 72–81. <https://doi.org/10.31102/jatim.v2i2.1274>
- Widiyaningtyas, T., Hidayah, I., & Adji, T. B. (2021). User profile correlation-based similarity (UPCSim) algorithm in movie recommendation system. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00425-x>
- Yoshua, I., Bunyamin, H., & Si, S. (2021). *Pengimplementasian Sistem Rekomendasi Musik Dengan Metode Collaborative Filtering (Vol. 3)*.