

PENERAPAN SISTEM REKOMENDASI PADA WEBSITE STREAMING MOVIE MENGUNAKAN METODE COLLABORATIVE FILTERING

Ani Isnaeni
Departemen Teknologi Informasi
Institute Perbanas
Jakarta, Indonesia
ani.isnaeni@perbanas.id

Valentinus Paramarta
Departemen Teknologi Informasi
Institut Perbanas
Jakarta, Indonesia
valentinus13@perbanas.id

Abstrak— Industry perfilman yang berkembang sangat pesat. Banyak sekali film yang beredar dan tersedia di berbagai platform streaming. Oleh karena itu banyak masyarakat yang kebingungan untuk memilih film apa yang bagus untuk ditonton dan platform streaming apa yang akan digunakan. Maka dibutuhkan sebuah sistem yang untuk merekomendasikan film yang menarik untuk ditonton dan sesuai dengan karakteristik pengguna streaming film maupun calon pengguna streaming film. penelitian ini bertujuan untuk membuat sistem rekomendasi berdasarkan rating dari pengguna platform streaming. Dalam membuat sistem rekomendasi tersebut, penelitian ini menggunakan metode collaborative filtering dan algoritma k-NN. Data film dan rating dari beberapa platform streaming (Netflix, HBO, Disney Hostar dan primeTV) yang telah dikumpulkan diolah dengan menggunakan algoritma KNNBaseline dan dilakukan pengujian dengan menggunakan nilai RMSE, precision dan recall. Hasil yang diperoleh penelitian ini adalah nilai RMSE yang didapatkan dari pengujian model sebesar 0,86, nilai precision terbaik adalah sebesar 0,91, dan nilai recall terbaik adalah 0,97 pada nilai k sebesar 30.

Kata Kunci— *platform streaming, kNN, sistem rekomendasi, collaborative filtering*

I. PENDAHULUAN

Perkembangan teknologi saat ini semakin cangih dan cepat. Hal ini dibuktikan dengan banyak nya inovasi

dibidang teknologi. Banyak mesin canggih yang diciptakan untuk mempermudah kehidupan manusia salah satunya adalah machine learning. Pada buku Pengantar Machine Learning yang ditulis oleh Bedy Purnama 2019 Machine learning merupakan sebuah machine yang terkait dengan komputer, machine learning bertujuan supaya komputer dapat belajar sendiri, berfikir sendiri, serta mengambil keputusan sendiri. Saat ini ada banyak sekali teknologi yang menerapkan machine learning. tanpa disadari, machine learning sangat dekat dengan kehidupan manusia. Salah satu implementasi machine learning yang sering ditemui di kehidupan sehari hari adalah sistem rekomendasi. Sistem rekomendasi ini dapat diterapkan diberbagai aplikasi seperti e-commerce, aplikasi pemutar music dan juga sistem rekomendasi pada platform streaming film.

Pada platform streaming film, sistem rekomendasi digunakan untuk mencari film yang menarik untuk ditonton oleh pengguna. Contoh dari sistem rekomendasi itu sendiri adalah ketika pengguna baru dari sebuah platform streaming film ingin mencari film horor yang bagus untuk ditonton, pengguna tersebut akan mencari film berdasarkan genre pada search engine yang ada pada platform streaming. Maka secara otomatis platform streaming tersebut akan menampilkan film horor dengan rating yang bagus yang telah ditonton banyak pengguna lainnya. Contoh lain dari sistem rekomendasi adalah, pada Platform E-commerce, biasanya pengguna akan mencari

barang yang ia cari di search engine platform E-commerce secara otomatis platform tersebut akan menampilkan barang yang mirip dengan apa yang dicari oleh pengguna. Sistem rekomendasi didefinisikan sebagai strategi pengambilan keputusan untuk pengguna di bawah lingkungan informasi yang kompleks [1]. Menurut jurnal yang ditulis oleh F.O Isinkaye dkk tahun 2015 sistem rekomendasi juga didefinisikan dari perspektif E-commerce sebagai alat yang membantu pengguna mencari melalui catatan pengetahuan yang terkait dengan minat dan preferensi pengguna. Sedangkan sistem rekomendasi secara umum adalah sistem yang dibuat untuk menentukan rekomendasi suatu item kepada penggunanya [2]. Sistem rekomendasi akan berpengaruh terhadap perilaku dan kebiasaan konsumen karena sistem ini akan menampilkan item paling diminati atau item yang terakhir kita lihat sehingga membuat konsumen merasa nyaman. Hal inilah yang membuat sistem rekomendasi sangat berguna bagi sebuah perusahaan khususnya platform digital seperti E-commerce dan platform penyedia streaming film.

Semakin berkembang dunia perfilman seperti sekarang banyak studio pembuatan film merilis film yang menarik untuk ditonton yang tersebar di berbagai platform streaming film. Hampir setiap website platform streaming memiliki judul film yang berbeda didalamnya. Contohnya film yang terdapat pada Netflix belum tentu tersedia di platform streaming lainnya. Bisa dikatakan bahwa setiap platform streaming film memiliki ciri khasnya masing – masing. Adanya perbedaan peminatan antara pengguna platform streaming dan juga perbedaan judul film yang tersedia pada platform streaming film inilah yang membuat pengguna membutuhkan referensi untuk memilih platform streaming mana yang akan digunakan untuk menonton film dan judul film apa yang akan ia tonton. Untuk membantu mencari referensi apa yang akan diberikan kepada pengguna, sistem rekomendasi akan sangat membantu pengguna untuk menemukan apa yang mereka inginkan. Sistem ini akan membantu dengan memunculkan film yang memiliki rating bagus pada beberapa platform streaming sehingga pengguna dapat mengetahui film apa yang bagus untuk ditonton sesuai peminatannya.

Untuk memberikan rekomendasi film yang bagus berdasarkan rating kepada pengguna pada Platform Streaming film, dapat menggunakan salah satu metode sistem rekomendasi yaitu collaborative filtering dimana metode ini bekerja dengan memberikan rekomendasi pada pengguna berdasarkan perilaku atau opini pengguna lainnya. Opini yang dimaksud disini adalah feedback dari pengguna lain seperti rating. Collaborative filtering merupakan proses penyaringan atau pengevaluasian item dengan menggunakan opini dari orang lain. Ide utamanya

adalah untuk mengeksploitasi informasi mengenai perilaku di masa lampau maupun opini dari suatu komunitas pengguna yang kemudian digunakan untuk memprediksi item mana yang akan disukai atau menarik bagi seorang pengguna [2]. Menurut artikel yang ditulis Ari Laksito pada tahun 2021 Cara kerja Teknik collaborative filtering ini adalah dengan memanfaatkan data pada komunitas dengan cara mencari kemiripan antar pengguna, yaitu mengasumsikan bahwa pengguna yang memiliki preferensi serupa di masa lalu cenderung memiliki preferensi yang sama di masa depan. Pada dasarnya kita akan lebih percaya dengan rekomendasi dari orang yang memiliki preferensi sama dengan kita, inilah dasar yang digunakan oleh collaborative filtering dalam generate item rekomendasi. Terdapat berbagai algoritma machine learning yang dapat digunakan pada metode collaborative filtering. salah satunya adalah algoritma k-NN. menurut penelitian yang dilakukan oleh Gede Pivin Suwirmayanti pada tahun 2017, Algoritma k-Nearest Neighbor (k-NN) ini dipilih untuk sistem rekomendasi karena metode k-NN merupakan suatu bentuk model pendukung keputusan yang dapat mengklasifikasikan data berdasarkan jarak terdekat dan k-NN dapat membantu menemukan pengguna dengan karakteristik serupa.

Berdasarkan permasalahan yang telah diuraikan diatas terkait dengan kesulitan pengguna platform streaming film untuk memilih film yang akan ditonton dari berbagai platform film yang ada serta berdasarkan dari beberapa penelitian terdahulu yang telah dilakukan maka peneliti pada penelitian ini menggunakan metode collaborative filtering menggunakan algoritma k-NN untuk merancang sistem rekomendasi film dari berbagai platform streaming yang ada (Netflix, HBO, Disney Hoster dan filmTV).

Rumusan masalah yang dapat disusun dalam penelitian ini adalah sebagai berikut:

1. Bagaimana metode Collaborative Filtering dan algoritma kNN dapat digunakan untuk merekomendasikan film yang diminati oleh pengguna berdasarkan dari berbagai platform streaming film?
2. Bagaimana performa metode Collaborative Filtering dengan menggunakan algoritma KNN dalam penerapan sistem rekomendasi berbagai platform streaming film?

II. LANDASAN TEORI

Pada bagian ini akan dijelaskan terkait penelitian terdahulu yang telah dilakukan. Penelitian yang dijelaskan pada bagian ini adalah penelitian yang memiliki topik terkait penggunaan berbagai algoritma *machine learning* dalam penelitian ini.

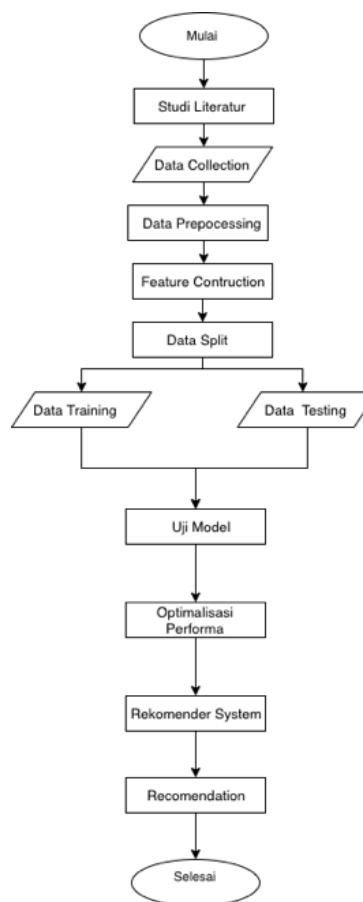
A. Penelitian Terdahulu Terkait Rekomendasi Pada Website Streaming Movie

Terdapat berbagai penelitian yang dilakukan oleh peneliti terdahulu terkait dengan sistem rekomendasi dengan menggunakan metode collaborative filtering dan algoritma k-NN. Pada tahun 2020 Eggy Ryana Agustian, melakukan penelitian terkait dengan Sistem Rekomendasi menggunakan arsitektur k-NN [3]. pada penelitian tersebut dijelaskan bahwa terdapat banyak film yang beredar, banyak penggemar menonton kesulitan untuk menentukan film apa yang akan disaksikan. Sehingga peneliti membuat sistem rekomendasi berbasis website dengan menggunakan metode collaborative filtering dan arsitektur k-Nearest Neighbor untuk memprediksi rating suatu film. Hasil dari penelitian ini adalah aplikasi dapat memberikan rekomendasi terhadap user aktif dengan cara menghitung rating prediksi dengan kondisi beberapa user lainnya sudah pernah memberi rating pada film yang akan diprediksi user aktif. Penelitian lain juga dilakukan oleh Ilham Gumantung Gusti, Dkk, pada tahun 2019 yang berjudul “Rekomendasi Pemilihan Mobil Menggunakan K-Nearest Neighbor Collaborative Filtering” tujuan penelitian ini adalah untuk memudahkan pengguna mobil dalam memilih jenis mobil yang diinginkannya [4]. Dalam penelitian yang dilakukan, peneliti menerapkan metode collaborative filtering dan arsitekturnya K-Nearest Neighbor yang dilakukan berdasarkan kedekatan jarak data testing dan data training. Kedekatan tersebut digunakan untuk merekomendasikan mobil ke pengguna. Adapun hasil yang diperoleh pada penelitian yang mereka lakukan adalah jika ingin mendapatkan 10 mobil terbaik maka jarak maksimal yang digunakan adalah 5%, dan akurasi terbaik didapatkan ketika $k=10$ yaitu sebesar 95,15%.

B. Algoritma Machine Learning

Penelitian yang dilakukan oleh Eggy Ryana Agustian, Munir, Eddy Prasetyo Nugroho (2020) memiliki tujuan untuk membuat sistem rekomendasi yang ditujukan pada pengguna mengingat banyak sekali film yang beredar. Metode yang digunakan pada penelitian yang dilakukan adalah metode collaborative filtering dengan menggunakan algoritma K-Nearest Neighbor. Hasil yang diperoleh yaitu aplikasi dapat memberikan rekomendasi terhadap user aktif dengan cara menghitung rating prediksi dengan kondisi beberapa user lainnya sudah pernah memberikan rating pada film yang akan diprediksi dan direkomendasikan untuk user aktif lainnya [5].

Penelitian yang dilakukan Ilham Gumantung Gusti, Muhammad Nasrum, Ratna Astuti Nugrahaeni (2019) memiliki untuk membantu customer untuk mengetahui informasi mobil yang sejenis dengan mobil yang disukainya. Cara kerja sistem rekomendasi pada penelitian



Gambar 1. Tahapan Penelitian

yang dilakukan adalah pengguna akan mendapatkan informasi lebih mengenai mobil yang ingin dipilih dan mobil lain yang mungkin belum diketahui sama sekali oleh customer tersebut. Metode yang digunakan pada penelitian ini adalah collaborative filtering dengan algoritma K-Nearest Neighbor hasil prediksi dihitung berdasarkan jarak kedekatan data testing dan data training. Kedekatan data (kemiripan data) tersebut digunakan untuk merekomendasikan mobil ke pengguna. Hasil yang diperoleh dalam penelitian ini adalah jika ingin mendapatkan 10 mobil terbaik maka jarak maksimal yang digunakan adalah 5%, dan akurasi terbaik didapatkan ketika $K = 10$ yaitu sebesar 95,15% [4].

III. METODOLOGI

Pada bagian ini akan dijelaskan data yang digunakan dalam penelitian, metode dan alat bantu yang digunakan, dan tahapan penelitian yang dilakukan.

A. Data Penelitian

Pada penelitian ini menggunakan data yang terdapat pada empat platform streaming yaitu Netflix, HBO, Disney Hostar, Prime TV. Data pada penelitian ini merupakan dataset yang memiliki jumlah 5021 baris 10

kolom yaitu userID, movieID, year, genre, Netflix, HBO, Disney Hoster, Prime TV dan Rating yang digunakan sebagai fitur. Adapun penjelasan terkait kolom yang terdapat pada dataset yang digunakan adalah sebagai berikut:

1. userID : merupakan kolom yang berisikan id user yang menonton film sekaligus memberikan rating pada film.
2. MovieID : merupakan kolom yang berisikan id dari film yang tersedia pada data.
3. Year : merupakan tahun dimana film di rilis.
4. Genre: merupakan genre dari film yang terdapat pada data.
5. Netflix, HBO, Disney Hoster dan Prime TV : merupakan platform streaming dimana film ditayangkan.
6. Rating : merupakan penilaian film dari user yang menonton.

B. Tahapan Penelitian

Dalam melakukan penelitian ini, dilakukan beberapa tahapan seperti berikut: pada Gambar 1.

a) Studi Literatur

Tahapan awal yang dilakukan dalam penelitian adalah melakukan studi literatur atau mengumpulkan beberapa jurnal terdahulu (5 tahun terakhir) yang berkaitan dengan penelitian yang sedang dilakukan. Studi literatur dilakukan untuk membandingkan metode dan hasil dari penelitian terdahulu dengan penelitian yang akan dilakukan. pada studi penelitian ini meliputi metode Collaborative filtering, k-NN, dan pengujian evaluasi pada penelitian terdahulu.

b) Crawling (Penarikan Data)

Tahap pengumpulan data merupakan hal terpenting dalam penelitian, dimana data adalah inti dari penelitian yang akan dilakukan. Data yang diperoleh peneliti merupakan dua data yaitu data film dan data rating keduanya merupakan data sekunder yang diperoleh dari situs Kaggle.

c) Pre-processing

Data yang diperoleh dari situs Kaggle, selanjutnya dilakukan pre-processing data. Dimana pre-processing data ini perlu dilakukan pada sebuah penelitian untuk memastikan bahwa kualitas data yang akan diteliti dalam keadaan baik dan bagus sebelum dilakukan analisis data serta penelitian lebih lanjut.

d) Feature Contruction:

Merupakan metode yang digunakan untuk menambahkan atribut pada data yang digunakan untuk penelitian. Feature Contruction yang dilakukan oleh peneliti adalah dengan melakukan manipulasi data dimana peneliti menambahkan kolom binary_platform dan kolom platform pada dataset.

e) Split Data

Data split pada penelitian ini merupakan tahapan dimana data dibagi menjadi 2 kategori yaitu data training dan data testing dengan perbandingan 25:75. Berikut ini merupakan pembahasan lebih lanjut mengenai data training dan data testing. Data training merupakan 75% dari keseluruhan data yang digunakan untuk melatih algoritma (training). Data testing merupakan 25% bagian dari data yang digunakan untuk menguji performa algoritma yang digunakan pada penelitian.

f) Uji Performa

Proses yang terjadi pada tahap uji model yaitu peneliti melakukan pengujian terhadap tiga algoritma k-NN yang digunakan untuk sistem rekomendasi yaitu KNNBaseline, KNNBasic dan KNNWithmean dari ketiga model tersebut yang memiliki hasil pengujian terbaik akan digunakan untuk mengolah data penelitian. Hasil dari pengujian model disini adalah nilai RMSE. Model yang memiliki nilai RMSE yang paling kecil merupakan model yang terbaik.

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Tahapan penting yang dilakukan pada penelitian ini yaitu pengumpulan data. Data yang digunakan merupakan data film yang diperoleh dari situs <https://www.kaggle.com/> . Untuk kebutuhan penelitian, peneliti melakukan modifikasi pada data yang telah diperoleh dari sumber data. Adapun modifikasi yang dilakukan oleh peneliti yakni menambahkan kolom userID untuk masing – masing judul film serta menambahkan kolom platform streaming yaitu Netflix, HBO, Disney Hoster, dan Prime TV. Untuk menemukan data userID peneliti mencari data sekunder lainnya pada situs Kaggle. Selanjutnya untuk setiap userID dilakukan penilaian berupa rating film untuk setiap judul film. Adapun penentuan rating dilakukan dengan cara menggabungkan dua data set yaitu dataset “ movies” dan dataset “ratings”.

B. Pre-processing Data

Setelah peneliti melakukan pengumpulan data, Langkah selanjutnya adalah preprocessing data. Tahap preprocessing data ini dilakukan untuk memastikan bahwa

data yang digunakan sudah dalam kondisi bagus dan sesuai kebutuhan untuk dilakukan analisis lebih lanjut. Pada tahap ini peneliti melakukan tahapan preprocessing yaitu feature construction dimana peneliti menambahkan atribut pada data asli. Peneliti menambahkan kolom `binary_platform` dan kolom `platform` pada dataset.

	Netflix	HBO	Disney Hstar	Prime TV	binary_platform	platform
0	1	1	0	1	1101	13
1	1	1	0	1	1101	13
2	1	1	0	1	1101	13
3	1	1	0	1	1101	13
4	1	1	0	1	1101	13
...	...	...	...	...	...	...
5915	0	0	0	1	0001	1
5916	1	0	0	1	1001	9
5917	0	0	0	1	0001	1
5918	0	0	1	1	0011	3
5919	0	0	0	1	0001	1

Gambar 2 dataset setelah dilakukan tahapan preprocessing data

Gambar 2 diatas merupakan data setelah dilakukan feature construction dimana kolom `binary_platform` digunakan untuk menandai platform streaming yang memiliki nilai 1 dan 0. Pada kolom binary ini data platform streaming akan digabungkan menjadi satu kesatuan sehingga menjadi bilangan biner. Sedangkan kolom `platform` merupakan bilangan asli yang diterjemahkan dari bilangan biner yang ada pada kolom `binary_platform`.

C. Uji Model

Tahapan selanjutnya adalah melakukan pengujian terhadap beberapa algoritma

k-NN. Algoritma yang dilakukan pengujian oleh peneliti adalah KNNBaseline, KNNBasic dan KNNWithmean. Pada tahap Pengujian model ini, peneliti menggunakan metode Cross Validation dari library surprise untuk menemukan nilai RMSE di setiap algoritma k-NN. Dari pengujian dengan menggunakan Cross Validation diperoleh hasil bahwa nilai RMSE dari KNNBaseline adalah yang paling kecil dibanding dengan yang lainnya.

Tabel 1. Tabel hasil pengujian performa beberapa algoritma k-NN

Algoritma	RMSE
KNNBaseline	0,9
KNNWithMeans	0,95
KNNBasic	1.0

Berdasarkan tabel 1 hasil pengujian model diatas, KNNBaseline dipilih untuk di implementasikan dalam pengolahan data pada penelitian. Hal ini terjadi karena semakin kecil nilai RMSE (mendekati nilai 0) menandakan bahwa semakin sedikit eror yang dihasilkan.

D. Optimalisasi Performa

Tahapan selanjutnya yang dilakukan oleh peneliti adalah optimalisasi performa dengan menggunakan GridsearchCV. Proses ini dilakukan untuk menemukan parameter terbaik dari algoritma yang digunakan. Gridsearch yang digunakan oleh peneliti adalah dengan menggunakan library surprise dengan rentang nilai k yang digunakan adalah 30 sampai 50 sedangkan rentang nilai `mean_k` nya yaitu antara 1 sampai 3. Adapun hasil yang dari proses ini yaitu nilai RMSE terbaik dan jumlah nilai k dan nilai mean k untuk algoritma yang digunakan.

Dari hasil proses optimalisasi performa, dengan menggunakan GridsearchCV ditemukan nilai RMSE terbaiknya yaitu 0,89 dan nilai k maksimum nya yaitu 30 sedangkan untuk nilai mean k nya yaitu diangka 1. Seperti pada gambar 4.4.

```
Done computing similarity matrix.
0.896997761195037
{'k': 30, 'mean_k': 1}
```

Gambar 3 Hasil dari proses optimalisasi performa dengan menggunakan gridsearchCV

Peneliti menggunakan hasil dari proses ini sebagai acuan untuk pengujian performa model algoritma yang digunakan.

E. Uji Performa

Langkah selanjutnya yang dilakukan peneliti adalah menguji algoritma yang digunakan yaitu KNNBaseline. pada proses pengujian kinerja algoritma, peneliti menggunakan `train_test_split()` untuk mengambil sampel data train dan data test dengan perbandingan ukuran 75:25 dimana 75% adalah ukuran data train dan 25% adalah data test. Kemudian, peneliti menggunakan metode `fit()` yang akan melatih algoritma dan metode `test()` yang akan mengembalikan prediksi yang dibuat dari testset.

Dari pengujian performa model yang telah dilakukan, peneliti mendapatkan hasil nilai RMSE sebesar 0,86.

```
Estimating biases using als...
Computing the msd similarity matrix...
Done computing similarity matrix.
RMSE: 0.8690
```

: 0.8690076269148383

Gambar 4 Hasil pengujian performa model KNNBaseline

Untuk melihat seberapa bagus model algoritma yang digunakan, peneliti melakukan pengujian dengan melihat rekomendasi film terbaik dan rekomendasi film terburuk.

best_predictions						
	uid	iid	rui	est	details	Ui err
1181	495	98154	5.0	5.000000	{'actual_k': 2, 'was_impossible': False}	2 0.000000
519	418	137337	4.0	4.000181	{'actual_k': 1, 'was_impossible': False}	1 0.000181
746	305	140110	4.0	4.000292	{'actual_k': 8, 'was_impossible': False}	9 0.000292
29	448	75805	2.0	2.000346	{'actual_k': 5, 'was_impossible': False}	5 0.000346
996	119	99112	4.0	3.999567	{'actual_k': 8, 'was_impossible': False}	8 0.000433
1435	528	80463	3.5	3.499554	{'actual_k': 32, 'was_impossible': False}	39 0.000446
308	68	81847	4.0	3.999107	{'actual_k': 19, 'was_impossible': False}	19 0.000893
156	17	80463	4.0	4.003881	{'actual_k': 30, 'was_impossible': False}	39 0.003881
457	318	137337	4.0	3.994113	{'actual_k': 1, 'was_impossible': False}	1 0.005887
477	596	176051	3.5	3.505897	{'was_impossible': False}	0 0.005897

Gambar 5 Rekomendasi film terbaik

worst_predictions						
	uid	iid	rui	est	details	Ui err
1064	461	104841	0.5	3.592140	{'actual_k': 17, 'was_impossible': False}	19 3.092140
167	63	91500	0.5	3.621654	{'actual_k': 35, 'was_impossible': False}	45 3.121654
1135	380	179819	0.5	3.730598	{'actual_k': 9, 'was_impossible': False}	9 3.230598
552	537	103341	0.5	3.824919	{'actual_k': 16, 'was_impossible': False}	16 3.324919
140	308	79132	1.0	4.415154	{'actual_k': 35, 'was_impossible': False}	110 3.415154
950	393	76251	0.5	3.950242	{'actual_k': 27, 'was_impossible': False}	31 3.450242
1334	258	122886	0.5	3.969053	{'actual_k': 18, 'was_impossible': False}	34 3.469053
1400	418	111781	0.5	3.977680	{'actual_k': 11, 'was_impossible': False}	11 3.477680
1304	418	89745	0.5	4.023845	{'actual_k': 35, 'was_impossible': False}	50 3.523845
620	495	106489	0.5	4.194669	{'actual_k': 17, 'was_impossible': False}	20 3.694669

Gambar 6 Rekomendasi film terburuk

Dari pengujian yang dilakukan didapatkan hasil bahwa semakin kecil selisih antara rating aktual dan rating yang di prediksi oleh sistem maka hasil prediksinya semakin baik. Sehingga, daftar movieid yang memiliki nilai err yang kecil masuk ke dalam kategori *best predictions*.

V. KESIMPULAN

Dalam penelitian yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut (1) Algoritma kNN berhasil diimplementasikan dalam sistem rekomendasi film yang ditunjukkan dari beberapa hasil pengujian pada user_id yang berbeda yang telah dilakukan ('user_id' = 50, 358 dan 274). (2) Nilai RMSE terbaik yang diperoleh dari proses hyperparameter tuning dengan menggunakan GridsearchCV adalah 0,8 dengan parameter 'k' yang didapatkan adalah 30 dan nilai parameter 'mean_k' adalah 1. Hasil dari proses ini digunakan sebagai acuan pengujian performa model.

REFERENSI

- [1] Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. (2015). Recommendation systems: Principles, Methods and Evaluation. *Egyptian Informatics Journal*, 261- 273.
- [2] Dzumiroh, L. (2012). Penerapan Metode Collaborative Filtering Menggunakan Rating Implisit pada Sistem Rekomendasi Pemilihan Film di Rental VCD.
- [3] Agustian, E. R., & Nugroho, E. P. Sistem Rekomendasi Film Menggunakan Metode Collaborative Filtering dan K-Nearest Neighbors. *Jurnal Aplikasi dan Teori Ilmu Komputer*, 3(1), 18-21.
- [4] Gusti, I. G., Nasrun, M., & Nugrahaeni, R. A. (2019). Rekomendasi Sistem Pemilihan Mobil Menggunakan K-Nearest Neighbor (KNN) Collaborative Filtering. *Teknika-Jurnal Penelitian Dan Pengembangan Telekomunikasi, Kendali, Komputer, Elektrik, Dan Elektronika*, 4(1), 26-32.
- [5] Agustian, E. R., & Nugroho, E. P. Sistem Rekomendasi Film Menggunakan Metode Collaborative Filtering dan K-Nearest Neighbors. *Jurnal Aplikasi dan Teori Ilmu Komputer*, 3(1), 18-21.