

SISTEM REKOMENDASI FILM MENGGUNAKAN METODE NEURAL COLLABORATIVE FILTERING

MOVIE RECOMMENDATION SYSTEM USING NEURAL COLLABORATIVE FILTERING

Daniel Ary Nugroho¹, Chairisni Lubis², Novario Jaya Perdana³

^{1,2,3}Universitas Tarumanagara
chairisnil@fti.untar.ac.id

ABSTRACT

Films are growing rapidly thanks to the increasing number of internet users. According to USC (University of Southern California), users easily get bored if they cannot find the movie they want within 1.5 minutes or within 90 seconds and may switch to other platforms, resulting in a decrease in active users. This problem can be overcome by using a recommendation system that can provide film recommendations based on user ratings. In this design, the recommendation system used is Neural Collaborative Filtering (NCF). Testing using 3,235 film titles, 610 users, and 87,812 ratings produced the best scenario with a recall value of 69.6% and Normalized Discounted Cumulative Gain (NDCG) with a value of 81.4%. From the test results, it can be concluded that 69.6% of films that should have been recommended were successfully recommended, and the recommendation position was similar to the actual position in 81.4% of cases.

Keywords: Collaborative Filtering, Film, Artificial Neural Network, Recommendation System.

ABSTRAK

Film berkembang pesat berkat meningkatnya jumlah pengguna *internet*. Dengan semakin banyaknya pengguna *internet*, layanan *streaming film* perlu membuat konten yang menarik agar pelanggan tetap terlibat dan menghabiskan lebih banyak waktu untuk menikmati penawaran mereka. Menurut USC (University of Southern California), pengguna mudah bosan jika tidak dapat menemukan *film* yang diinginkan dalam waktu 1,5 menit atau dalam 90 detik dan dapat beralih ke *platform* lain, yang mengakibatkan penurunan pengguna aktif. Masalah ini dapat diatasi dengan menggunakan sistem rekomendasi yang dapat memberikan rekomendasi film berdasarkan rating pengguna. Pada perancangan ini, sistem rekomendasi yang digunakan adalah Neural Collaborative Filtering (NCF). Pengujian menggunakan 3.235 judul film, 610 pengguna, dan 87.812 *rating* menghasilkan skenario terbaik dengan nilai *recall* 69,6% dan *Normalized Discounted Cumulative Gain (NDCG)* dengan nilai 81,4%. Dari hasil pengujian tersebut dapat disimpulkan bahwa 69,6% *film* yang seharusnya direkomendasikan berhasil direkomendasikan, dan posisi rekomendasinya mirip dengan posisi sebenarnya pada 81,4% kasus.

Kata Kunci: Penyaringan Kolaboratif, Film, Jaringan Syaraf Tiruan, Sistem Rekomendasi

PENDAHULUAN

Film merupakan salah satu hiburan yang paling sering dinikmati di waktu senggang, para pecinta film umumnya menggunakan layanan penyedia streaming agar dapat mengakses lebih banyak variasi *film*. Saat ini, layanan streaming saling bersaing untuk menarik pengguna menggunakan produknya. Salah satu upaya yang dilakukan untuk menarik minat pengguna adalah dengan menggunakan sistem rekomendasi yang memungkinkan pengguna untuk mengakses *film-film* yang sulit dijangkau

jika mencari satu per satu, misalnya pengguna akan diberikan rekomendasi film sesuai riwayat interaksinya dengan film, seperti rekomendasi *film* bergenre horor jika pengguna sering menonton *film* bergenre tersebut, rekomendasi pengguna dapat dipengaruhi oleh kesamaan preferensi pengguna lain.

Seiring perkembangan pengguna internet, layanan *streaming film* dituntut untuk membuat konten berkualitas tinggi untuk menarik dan mempertahankan pengguna menggunakan layanan mereka lebih lama. Konten yang dihasilkan

memiliki variasi yang tinggi, membuat pengguna bingung dan mudah bosan saat memilih film, ingin menonton lalu memilih layanan streaming lainnya, dengan permasalahan tersebut muncul sistem rekomendasi yang mampu memberikan pilihan lebih cepat dan mengurangi jumlah pengguna yang bosan saat memilih film lebih dari 90 detik (Wang 2022).

Pada topik serupa dengan judul “Sistem Rekomendasi Film Berbasis Website Dengan Metode Prototype Menggunakan Algoritma K-Nearest Neighbors” oleh Alifia Rahma Fitrianti, Asih Rohman, Widjanarto memperoleh hasil rekomendasi dengan nilai presisi 45% dan recall 60% hanya dengan menggunakan atribut genre film. Penelitian ini berfokus pada pengembangan sistem yang mampu memberikan rekomendasi dengan menggunakan rating masing-masing pengguna, sehingga pengguna yang memiliki rating serupa akan saling merekomendasikan (Fitrianti, Rohmani, and Widjanarto 2020).

Collaborative filtering research umumnya menggunakan pengguna sebagai acuan rekomendasi, tetapi bisa juga berdasarkan *item*. Pradhan et al. (2021) menggunakan penyaringan kolaboratif berbasis item dan kesamaan kosinus untuk menghitung kesamaan antar fitur *item*, kemudian merekomendasikan kesamaan tertinggi kepada pengguna. Dalam studi yang sama menggunakan dataset MovieLens, Muhammad et al. (2022). menggunakan pemfilteran kolaboratif berbasis item yang menggabungkan K-Nearest Neighbors dengan formula kesamaan untuk memprediksi peringkat *film* yang direkomendasikan.

Sistem rekomendasi modern seringkali memanfaatkan Neural Networks sebagai salah satu komponen dalam arsitekturnya. Neural Networks banyak digunakan karena kemampuannya untuk mengenali pola dan hubungan yang kompleks pada pembelajaran mesin skala

besar. *Neural Collaborative Filtering* menggunakan jaringan saraf dengan tujuan menambahkan kernel non-linier ke faktorisasi matriks. (Priya et al. 2022) menggunakan pemfilteran berbasis konten dan jaringan saraf untuk merekomendasikan *video* di *YouTube*, di mana kedua komponen tersebut menghasilkan rekomendasi yang berpotongan untuk menghasilkan hasil rekomendasi akhir (Covington et al. 2016). menggunakan jaringan saraf yang lebih kompleks yang menggabungkan fitur seperti usia, bahasa, bahasa *video*, dan riwayat tontonan untuk memberikan peringkat *video* berurutan sebagai rekomendasi.

Pemfilteran berbasis konten, yang umumnya menggunakan fitur item, biasanya menghitung kesamaan menggunakan kesamaan kosinus atau kesamaan Pearson (Erlsha and Hiryanto 2017). menggunakan dua algoritma untuk merekomendasikan kursus yang dapat diambil. Algoritma pertama menghitung frekuensi pendaftaran kursus, sedangkan algoritma kedua, Naive Bayes, memperkirakan keberhasilan rekomendasi. (Fajriansyah et al. 2021). menggunakan metode untuk menghitung bobot judul film dan sinopsis. Setiap bobot kemudian digunakan untuk menghitung kesamaan menggunakan kesamaan kosinus. Langkah terakhir melibatkan pemfilteran berdasarkan *genre film*. Selain *film*, pemfilteran berbasis konten juga cocok untuk pengelompokan dokumen. (Parwita, Swari, and Welda 2018). melakukan *preprocessing* terhadap isi dokumen, seperti ekstraksi kata kunci, *tokenization*, *stopword removal*, *stemming*, pembobotan kalimat, dan terakhir menghitung kemiripan bobot dengan dokumen lain.

Sistem rekomendasi juga dapat memanfaatkan survei pengguna untuk memilih film, dimana pengguna diberikan pertanyaan untuk menentukan hasil rekomendasi. Beberapa pertanyaannya antara lain *genre* favorit, aktor favorit, tahun rilis, pentingnya rating di platform

seperti IMDB atau Rotten Tomatoes. Misalnya, jika pengguna memilih horor sebagai *genre* favoritnya, sistem rekomendasi kemungkinan akan menyediakan film horor. Metode ini dapat diimplementasikan seperti yang ditunjukkan oleh (Sandrya et al. 2022).

Pemfilteran kolaboratif juga dapat memanfaatkan data terstruktur seperti grafik untuk memetakan hubungan antara pengguna dan item. (Wibowo and Munir 2013) menggunakan dua grafik: grafik pengguna dan grafik item. Kesamaan cosinus dihitung dari grafik pengguna, dan pengguna dengan kesamaan tertinggi dikelompokkan menggunakan K-Nearest Neighbors dan direkomendasikan satu sama lain.

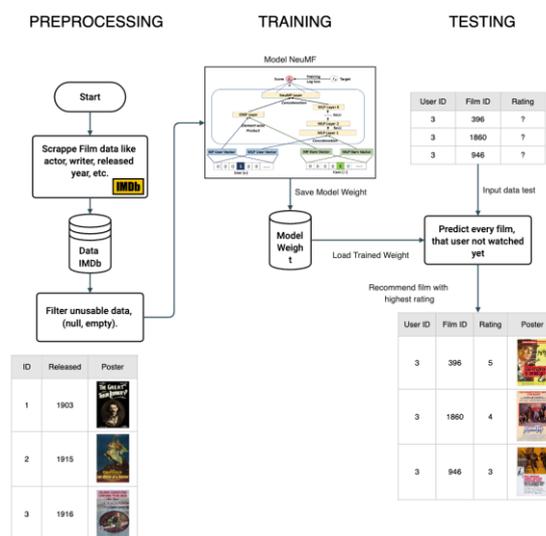
Tujuan dari penelitian ini adalah untuk membuat sebuah sistem rekomendasi untuk memprediksi rating pengguna yang tidak berinteraksi dengan sebuah film. Peringkat pada film mewakili minat atau kesukaan pengguna, semakin tinggi peringkat mensimulasikan minat pengguna terhadap film tersebut. Metode yang digunakan dalam perancangan ini adalah *Neural Collaborative Filtering (NCF)* yang dikembangkan dari algoritma *Matrix Factorization (MF)*, tujuan pengembangan adalah menambahkan kernel non-linear dari *Neural Network* untuk mengenali pola yang tidak dapat dikenali oleh *Matrix Factorization*.

METODE

Perancangan sistem rekomendasi ini menggunakan metode *Neural Collaborative Filtering* untuk memprediksi rating yang diberikan oleh pengguna terhadap film tersebut, film dengan rating tertinggi akan direkomendasikan terlebih dahulu, sedangkan film dengan prediksi rating di bawah tiga tidak muncul yang berarti pengguna tidak menyukai film tersebut. Model ini didasarkan pada *Neural Collaborative Filtering*, yaitu faktorisasi matriks yang menggunakan matriks jarang untuk merepresentasikan interaksi antara

pengguna dan film, dan mengisi *rating* kosong untuk film yang belum diberi *rating* oleh pengguna (Schafer and Frankowski 2007).

Model perlu dilatih, sebelum dapat digunakan. Model akan memprediksi berapa banyak peringkat yang akan diberikan dari pengguna ke film. Pada Gambar 1 merupakan alur kerja sistem sebagai berikut:



Gambar 1. Flowchart Diagram Pelatihan Model

Dataset

Perancangan ini menggunakan dataset yang disediakan oleh MovieLens, MovieLens merupakan komunitas virtual yang merekomendasikan film untuk ditonton oleh pengguna. *Dataset* tersebut berisi kumpulan rating pada *film* untuk membangun sistem rekomendasi, dataset tersebut dikumpulkan dan dikelola oleh sebuah organisasi bernama GroupLens dari University of Minnesota (Harper and Konstan 2015).

Data yang digunakan dalam MovieLens adalah data film, pengguna, dan rating, yang masing-masing berjumlah 9.280 film, 794 pengguna, dan 25 juta rating. Data perlu diproses sebelum dapat digunakan, seperti menghapus data yang tidak lengkap, memiliki kurang dari tiga pemain, tidak memiliki poster, atau tidak ada pengguna yang menilai film dalam data peringkat. *Filter* diterapkan pada data,

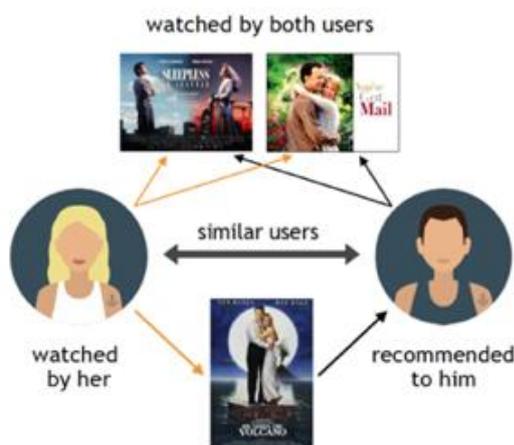
sehingga mengurangnya dari jumlah sebenarnya (Rapaport 2022).

Data yang disaring berjumlah 3.235 film, 610 pengguna, dan 87.813 rating. Data film memiliki atribut seperti id, nama film, tahun rilis, genre, poster. Data pengguna hanya memiliki atribut id. Data rating memiliki user id, film id, rating, dan timestamp sebagai tanggal rating.

Matrix Factorization

Matrix Factorization (MF) adalah algoritma yang dikembangkan pada kompetisi Netflix Prize 2009. MF merupakan jenis algoritma sistem rekomendasi collaborative filtering, artinya rekomendasi ditentukan berdasarkan kesamaan interaksi pengguna antara pengguna lain pada produk, dalam hal ini produk tersebut adalah film (Koren, Bell, and Volinsky 2009).

Collaborative Filtering



Gambar 2. Collaborative Filtering

Tujuan dari algoritma ini adalah untuk memprediksi rating yang akan diberikan pengguna terhadap produk, prediksi tersebut menggunakan data histori interaksi pengguna, seperti rating sebelumnya (Nvidia).

	FILM 1	FILM 2	FILM 3
USER 1	5	3	
USER 2	5	3	4
USER 3	5		

Gambar 3. Contoh Collaborative Filtering

Matrix Factorization menggunakan matriks *sparse* untuk membentuk interaksi antara pengguna dan film. MF bekerja dengan mendekomposisi matriks jarang menjadi dua matriks, satu untuk matriks pengguna p dan satu untuk matriks film q , masing-masing menyimpan satu set fitur laten atau karakteristik pengguna dan film. Tujuan dari faktorisasi matriks, adalah mengisi semua nilai pada tabel kosong dengan menggunakan data interaksi yang ada.

Operasi yang dilakukan oleh faktorisasi matriks adalah perkalian matriks sederhana yang menghasilkan matriks jarang beserta nilai prediksinya, kemudian menghitung *error*, dan memperbarui bobot masing-masing matriks fitur laten p dan q . Berikut adalah rumus faktorisasi matriks:

$$\hat{r}_{ui} = p_u q_i^T = \sum_{k=1}^K p_{uk} q_{ik} \quad (1)$$

Keterangan:

\hat{r} = nilai prediksi yang diberikan pengguna ke produk.

q = matriks fitur tersembunyi untuk produk.

p = matriks fitur tersembunyi untuk pengguna.

K = ukuran fitur tersembunyi.

Perhitungan *error* dilakukan untuk menghitung seberapa besar selisih nilai prediksi dengan nilai sebenarnya. Salah satu rumus yang paling umum digunakan untuk menghitung *error* adalah Mean Absolute Error dengan rumus di bawah ini:

$$MAE = \frac{1}{|\hat{R}|} \sum_{r_{ui} \in \hat{R}} |r_{ui} - \hat{r}_{ui}| \quad (2)$$

Keterangan:

R = ukuran data.

r = nilai Sebenarnya.

\hat{r} = nilai yang diprediksi.

Langkah terakhir dalam faktorisasi matriks adalah memperbarui fitur laten menggunakan dua rumus, rumus pertama memperbarui nilai fitur laten pada matriks pengguna dan formula kedua memperbarui nilai fitur laten pada matriks produk (Shahawy):

$$p_u \leftarrow p_u + \gamma \cdot (e_{ui} \cdot q_i - \lambda \cdot p_u) \quad (3)$$

$$q_i \leftarrow q_i + \gamma \cdot (e_{ui} \cdot p_u - \lambda \cdot q_i) \quad (4)$$

Keterangan:

p = fitur laten pengguna.

q = fitur laten produk.

e = nilai *error* prediksi.

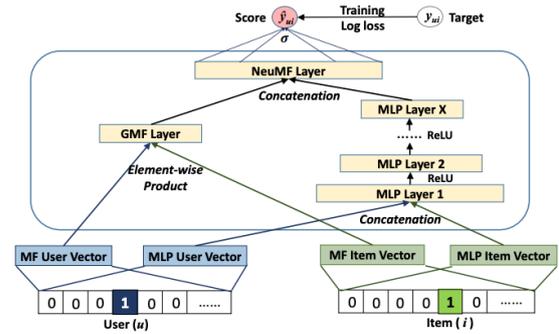
γ = *learning rate*.

λ = nilai regularizer.

Operasi faktorisasi matriks perlu dilakukan berulang kali mulai dari prediksi sampai dengan memperbarui nilai fitur laten, pengulangan akan berhenti ketika nilai *error* mencapai titik rendah yang ditentukan (Chen 2020).

Neural Collaborative Filtering

Neural Collaborative Filtering (NCF) atau Neural Matrix Factorization (NMF) adalah sebuah framework yang dibuat pada tahun 2017, NCF merupakan pengembangan dari *Matrix Factorization* dengan menambahkan komponen *Neural Network* yang bertujuan untuk mengenali pola interaksi pengguna dengan produk secara non linier, pola tersebut tidak dapat dikenali oleh *Matrix Factorization* pada umumnya. NCF menggunakan lapisan *embedding* untuk memetakan data pengguna dan produk, yang digunakan sebagai input. Gambar 4 di bawah ini adalah arsitektur dari model Neural Collaborative Filtering:



Gambar 4. Arsitektur Neural Collaborative Filtering

Neural Collaborative Filtering (NCF) memiliki dua komponen dalam arsitekturnya yaitu *Multilayer Perceptron* dan *Generalized Matrix Factorization* yang masing-masing berguna untuk mengenali pola interaksi dan mengekstrak fitur non-linear antara pengguna dan film. *Model NCF* menggunakan arsitektur yang rumit yang bertujuan untuk mengenali pola interaksi yang kompleks, yang bahkan tidak dapat dikenali oleh faktorisasi matriks (He, Liao, and Zhang 2020).

NCF memanfaatkan lapisan *embedding* yang berguna untuk menyimpan karakteristik laten atau fitur pengguna dan film. Komponen pertama dalam NCF adalah *Generalized Matrix Factorization* (GMF), GMF merupakan implementasi dari matriks faktorisasi dengan menambahkan kernel non linier seperti fungsi aktivasi. Rumus untuk GMF adalah sebagai berikut:

$$\hat{y}_{ui} = a_{out}(h^T(p_u \odot q_i)), \quad (5)$$

Keterangan:

p = fitur laten pengguna.

q = fitur laten produk.

a_{out} = aktivasi non linear.

h^T = bobot akhir pada *output layer*.

Komponen kedua dalam NCF adalah *Multilayer Perceptron* (MLP), MLP dalam NCF berguna untuk mengekstrak fitur interaksi yang kompleks antara pengguna dengan pengguna lain dan film. MLP dirumuskan sebagai berikut:

$$z_1 = \phi(p_u, q_i) = \begin{bmatrix} p_u \\ q_i \end{bmatrix},$$

$$\phi_2(z_1) = a_2(W_2^T z_1 + b_2),$$

$$\begin{aligned} & \dots\dots \\ \phi_L(z_{L-1}) &= a_L(W_L^T z_{L-1} + b_L), \\ \widehat{y}_{ui} &= \sigma(h^T \phi_L(z_{L-1})), \end{aligned} \tag{6}$$

Keterangan:

- p = fitur laten pengguna.
- q = fitur laten produk.
- a = aktivasi non linear.
- b = nilai bias pada layer L.
- h^T = bobot akhir pada output layer.
- W = nilai bobot untuk layer L dan posisi T.
- z = hidden layer untuk layer L.
- σ = fungsi aktivasi akhir.

Menggabungkan *MLP* dan *GMF* menjadi *model NCF*. Dengan menggabungkan kedua model tersebut, maka arsitektur *NCF* dapat menerima empat masukan yaitu dua masukan untuk *GMF* dan dua masukan untuk *MLP*. Berikut rumus penggabungan *GMF* dan *MLP*:

$$\begin{aligned} \phi^{GMF} &= p_u^G \odot q_i^G, \\ \phi^{MLP} &= a_L \left(W_L^T \left(a_{L-1} \left(\dots a_2 \left(W_2^T \begin{bmatrix} p_u^M \\ q_i^M \end{bmatrix} \right. \right. \right. \right. \right. \\ & \quad \left. \left. \left. \left. \left. + b_2 \right) \dots \right) \right) \right) + b_L \right), \\ \widehat{y}_{ui} &= \sigma \left(h^T \begin{bmatrix} \phi^{GMF} \\ \phi^{MLP} \end{bmatrix} \right), \end{aligned} \tag{7}$$

Keterangan:

- p^G = embedding pengguna komponen GMF.
- q^G = embedding produk komponen GMF.
- p^M = embedding pengguna komponen MLP.
- q^M = embedding produk komponen GMF.
- W^T = bobot MLP layer ke T.
- b = nilai bias pada MLP.
- h = bobot akhir pada output layer.
- σ = fungsi aktivasi MLP.
- a = fungsi aktivasi neural network.
- u = id pengguna.
- i = id produk.

Pengujian Sistem Rekomendasi

Pengukuran kualitas sistem rekomendasi umumnya menggunakan beberapa metrik seperti *Hit Ratio* (HR) atau *Recall* untuk menghitung persentase berapa *film* yang muncul dari data pengujian, *Normalized Discounted*

Cumulative Gain (*NDCG*) untuk menghitung efektivitas sistem ranking atau posisi produk yang relevan. Kedua metrik ini adalah beberapa metrik yang paling umum digunakan untuk mengukur kualitas sistem rekomendasi. Metrik umumnya menggunakan pengukuran Top-L atau top L film yang akan dijadikan sampel hasil rekomendasi terbanyak, nilai L yang biasa digunakan adalah lima dan 10. Berikut adalah rumus *Hit Ratio* dan *Normalized Discounted Cumulative Gain* (Benjamin 2021):

$$HR = \frac{|U_{hit}^L|}{|U_{all}|} \tag{8}$$

Keterangan:

- U_{hit}^L = Jumlah film yang berhasil direkomendasikan untuk setiap pengguna.
- U_{all} = Jumlah film yang harus direkomendasikan.

Rumus di atas dapat disimpulkan bahwa metrik *Hit Ratio* akan meningkat dengan jumlah L yang digunakan, karena kemungkinan besar jawaban yang benar masuk dalam daftar rekomendasi (Alsini 2020).

Untuk metrik *NDCG*, perlu ditentukan nilai relevansi antara film yang direkomendasikan dengan film yang seharusnya direkomendasikan, nilai tersebut mampu memberikan makna yang lebih detail terhadap hasil rekomendasi. Berikut adalah bagaimana nilai relevansi didefinisikan (Dhinakaran 2023).

Jika posisi film yang direkomendasikan sama persis dengan posisi sebenarnya dari film yang direkomendasikan, skor relevansinya adalah tiga (3). Jika posisi *film* yang direkomendasikan tepat di depan atau di belakang posisi sebenarnya dari daftar rekomendasi, maka nilai relevansinya adalah dua (2). Jika film yang direkomendasikan jauh dari posisi sebenarnya dari daftar rekomendasi, maka nilai relevansinya adalah satu (1). Jika film yang seharusnya direkomendasikan, tidak muncul dalam daftar film direkomendasikan teratas, nilai relevansinya adalah nol (0).

Nilai relevansi yang telah ditentukan digunakan untuk menghitung nilai *DCG* dan *IDCG*. Berikut adalah rumus *DCG*:

$$DCG = \sum_{i=1}^{ranks} \frac{gains}{\log_2(i + 1)} \quad (9)$$

Keterangan:

gains = relevansi hasil rekomendasi pada posisi *i*.
ranks = jumlah film yang direkomendasikan.

$$NDCG = \frac{DCG}{IDCG} \quad (10)$$

Keterangan:

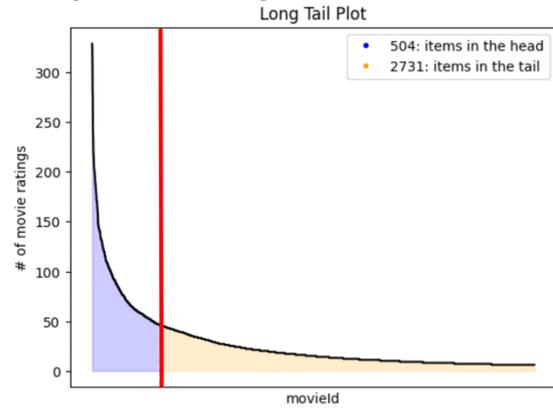
DCG = relevansi untuk setiap hasil rekomendasi.
IDCG = urutan nilai relevansi terbaik berdasarkan urutan sebenarnya.

HASIL DAN PEMBAHASAN

Pengujian *Neural Collaborative Filtering (NCF)* akan berupa pengujian *Generalized Matrix Factorization*, pengujian *Multilayer Perceptron*, dan pengujian kedua komponen yang digabungkan menjadi satu yang disebut *NCF*, tujuan pengujian ketiga *model* tersebut adalah untuk mengukur kinerja komponen dalam *NCF* yaitu *GMF* dan *MLP* berdasarkan penurunan nilai *loss* selama pelatihan, semakin kecil nilai *loss* menunjukkan *model* tersebut layak untuk pembelajaran data, selain itu untuk membuktikan apakah *model GMF* dan *MLP* layak digabungkan menjadi *NCF* untuk meningkatkan kualitas data. hasil rekomendasi.

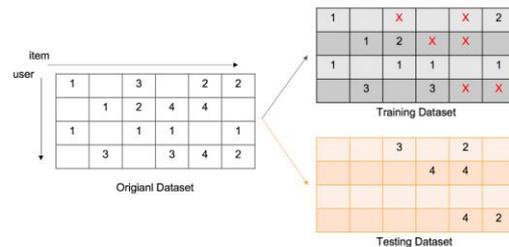
Dataset dieksplorasi terlebih dahulu untuk mendapatkan pengetahuan data yang lebih dalam, salah satu metode *plot* yang digunakan adalah long tail plot untuk menggambarkan keseimbangan rating setiap film. Gambar 5 merupakan plot yang menampilkan 50% dari total rating berada pada 504 film pertama, sedangkan 50% lainnya terbagi atas 2.731 film sisanya, dapat disimpulkan bahwa setiap film tidak memiliki jumlah rating yang seimbang, karena setengah dari data berada pada porsi kecil film tersebut. Data yang tidak seimbang dapat disebabkan oleh kurangnya interaksi antara pengguna

dengan produk, yang dapat mempengaruhi kinerja model (Longo 2018).



Gambar 5. Long Tail Plot

Dataset yang digunakan untuk pengujian berasal dari *dataset* yang sama dengan data yang digunakan untuk pengujian. Pembagian *dataset* untuk *testing* dan *training* adalah 30% dan 70%. Pembagian data diambil secara acak (Browniee 2020).



Gambar 6. Pembagian Dataset

Ketiga *model* menggunakan *parameter* yang sama untuk *learning rate*, *batch size*, dan *epoch* yaitu 0.001, 32, dan 20. *Parameter* yang diuji dan diubah adalah *L* dari Top-L dan ukuran *embedding* pada *model*, *parameter* tersebut digunakan untuk mencari nilai *L* dan ukuran *embedding* yang paling optimal, untuk dibandingkan dengan kedua *model* lainnya. *k* untuk membandingkan dengan dua *model* lainnya.

Tabel 1. Parameter Pengujian

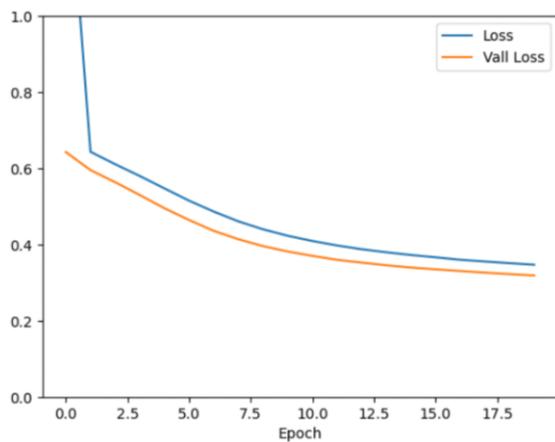
No	L	Ukuran Embedding
1	5	16
2	5	32
3	10	16
4	10	32

Pengujian *model GMF* menggunakan 20 epoch dengan 0,001 dan keempat parameter uji, menghasilkan nilai recall dan recall. NDCG pada Tabel 2

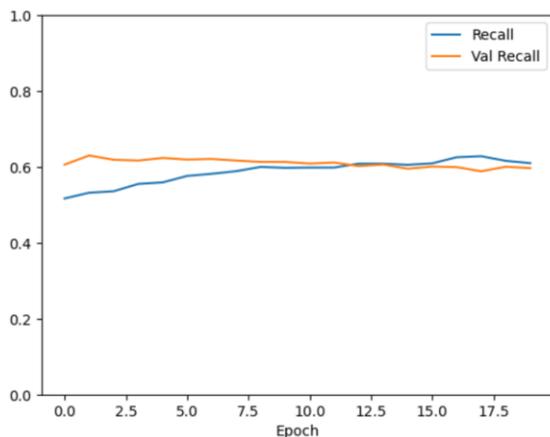
Tabel 2. Parameter Pengujian GMF

No	L	k	Recall	NDCG
1	5	16	0.607	0.798
2	5	32	0.588	0.773
3	10	16	0.605	0.772
4	10	32	0.596	0.769

Hasil pengujian model GMF menunjukkan bahwa Top-L dengan lima *film* teratas dan nilai prediksi 16 *film* bisa mendapatkan nilai *recall* di atas 60%. k sebanyak 16 buah diprediksi mampu mendapatkan nilai recall diatas 60%, dapat disimpulkan bahwa nilai paling optimal untuk model GMF adalah 16. k Nilai NDCG yang didapatkan cukup konsisten sebesar 77% untuk ketiga skenario, namun mencapai nilai terbaiknya pada parameter 5 untuk L dan 16 untuk ukuran *embedding* k.



Gambar 7. Loss Pelatihan GMF Dengan Parameter Terbaik



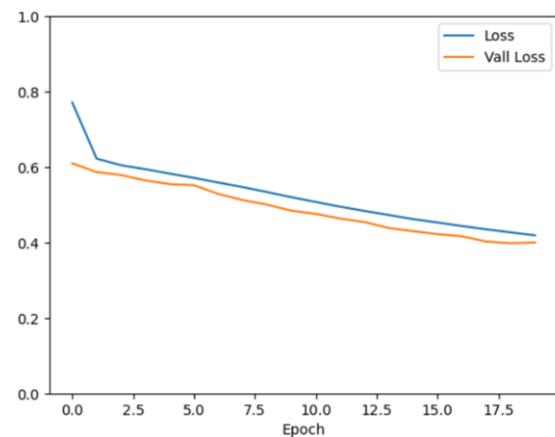
Gambar 8. Recall Pelatihan GMF Dengan Parameter Terbaik

Tabel 3. Parameter Pengujian MLP

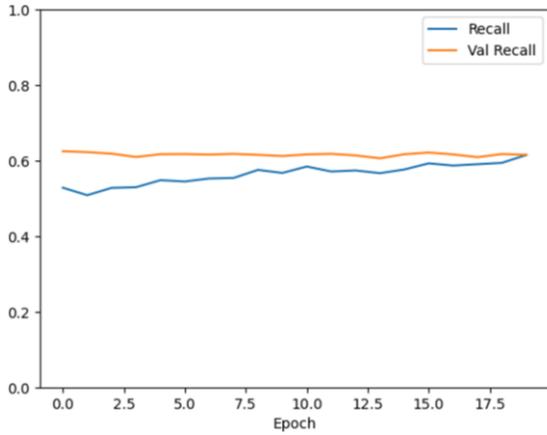
No	L	k	Recall	NDCG
1	5	16	0.598	0.795
2	5	32	0.614	0.770
3	10	16	0.613	0.771
4	10	32	0.608	0.763

Hasil pengujian model MLP menunjukkan bahwa dengan pengaturan Top-L lima film teratas dan nilai prediksi sebanyak 32, model ini mampu mencapai nilai recall di atas 60%. Dengan demikian, dapat disimpulkan bahwa nilai k yang optimal untuk model MLP adalah 32. Nilai NDCG juga menunjukkan konsistensi yang cukup baik, dengan persentase sebesar 77% untuk ketiga skenario yang diuji. Namun, parameter-parameter terbaik untuk model MLP adalah L sebesar 5 dan ukuran *embedding* k sebesar 16, di mana NDCG mencapai nilai terbaiknya.

Dengan pengaturan parameter ini, model MLP mencapai kinerja terbaik dalam menghasilkan rekomendasi film, dengan recall di atas 60% dan NDCG mencapai nilai tertinggi. Ini menunjukkan bahwa pengaturan parameter tersebut sangat penting dalam memaksimalkan kualitas rekomendasi yang diberikan oleh model.



Gambar 9. Loss Pelatihan MLP Dengan Parameter Terbaik

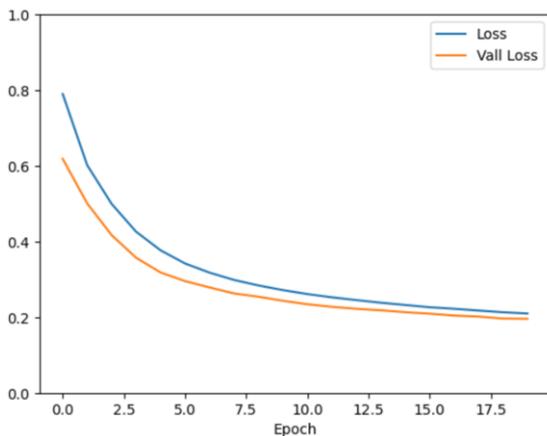


Gambar 10. Recall Pelatihan MLP Dengan Parameter Terbaik

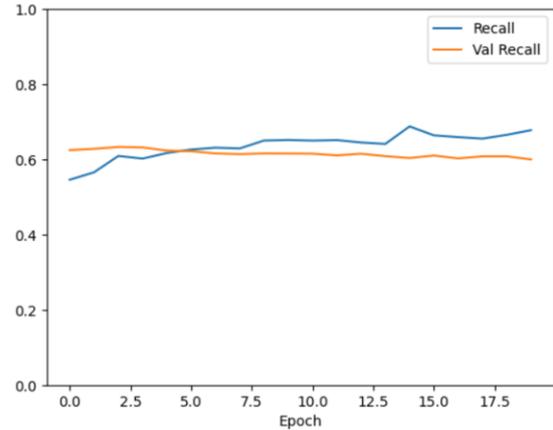
Tabel 4. Parameter Pengujian NCF

No	L	k	Recall	NDCG
1	5	16	0.625	0.816
2	5	32	0.696	0.814
3	10	16	0.624	0.778
4	10	32	0.686	0.767

Hasil pengujian model NCF menunjukkan bahwa Top-L dengan lima film teratas dan nilai prediksi sebanyak 32 buah mampu mendapatkan nilai metrik 69.6%. Dengan nilai k sebanyak 32 buah diprediksi mampu mendapatkan nilai metrik mendekati 70%, dapat disimpulkan bahwa nilai paling optimal untuk model NCF adalah 32. k Nilai NDCG yang didapatkan cukup konsisten sebesar 77% untuk ketiga skenario, namun mencapai nilai terbaiknya pada parameter 5 untuk L dan 16 untuk ukuran embedding k.



Gambar 11. Loss Pelatihan NCF Dengan Parameter Terbaik



Gambar 12. Recall Pelatihan NCF Dengan Parameter Terbaik

Setiap skenario memprediksi daftar film yang direkomendasikan berdasarkan peringkat film yang diprediksi pengguna. Berikut adalah tabel 5 contoh hasil rekomendasi terhadap model NCF, serta nilai rating aktual.

Tabel 5. Contoh Hasil Prediksi Rating

User ID	Item ID	k	Prediction
23	1447	3.455	3
16	1063	4.212	4.5
24	2672	5.205	5
51	2873	5.071	5
22	1147	4.065	4
7	1150	2.335	3
4	1330	4.041	4
23	1447	3.455	3

Setelah ketiga model diuji, dilakukan perbandingan model mana yang memiliki nilai loss dan recall terbaik diantara ketiganya, dan seberapa besar pengaruh kedua model tersebut terhadap model utama yaitu NCF.

Dari tabel pengujian dengan keempat skenario di atas, dipilih satu skenario dengan nilai recall terbaik pada masing-masing model.

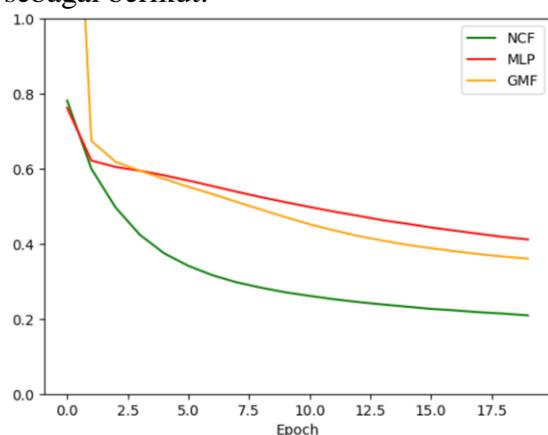
Tabel 6. Skenario Model Terbaik

Model	k	Recall	NDCG
NCF	32	0.696	0.814
MLP	32	0.614	0.770
GMF	16	0.607	0.798
Model	k	Recall	NDCG

Tabel 6 membandingkan metrik antara ketiga model, skenario terbaik untuk pengujian yang telah dilakukan adalah menggunakan nilai Top-L lima, NCF dan MLP menggunakan nilai k yang sama

yaitu 32, sedangkan *GMF* menggunakan nilai 16.k nilai *recall* dan *NDCG* pada *NCF* paling tinggi dibandingkan kedua model lainnya, persentase kenaikan *recall* lebih dari 10%.

Jika ketiga model tersebut dibandingkan nilai *loss* selama pelatihan maka grafik perbandingannya adalah sebagai berikut:



Gambar 13. Perbandingan Ketiga Model

Dalam gambar ke-10, terlihat perbandingan nilai *loss* selama pelatihan ketiga *model*. Jika diperhatikan, *model MLP* menghasilkan nilai *loss* paling tinggi dibandingkan dengan dua *model* lainnya. Hal ini mengindikasikan bahwa *model MLP* mungkin kurang cocok untuk mengenali pola interaksi antara pengguna dan produk (film).

Model GMF, meskipun memiliki nilai *loss* yang hampir mirip dengan *MLP*, menunjukkan sedikit peningkatan pada *epoch* 10 hingga 20. Sementara itu, *Model NCF* memiliki penurunan nilai *loss* yang konsisten sejak *epoch* ketiga hingga 20. Ini menunjukkan bahwa *Model NCF* memiliki kemampuan yang baik untuk mengurangi *loss* selama pelatihan. Lebih lanjut, jika dilatih dengan lebih banyak *epoch*, *model NCF* dapat menghasilkan nilai *loss* yang lebih kecil.

Dalam skenario terbaik, *Model NCF* membuktikan bahwa penggabungan *MLP* dan *GMF* merupakan pilihan yang tepat sebagai model prediksi rating pengguna terhadap produk (film). Ini terlihat dari peningkatan nilai *recall* pada model *NCF*. *MLP* dalam *Model NCF* berperan dalam

mengekstraksi fitur antara pengguna dan produk, dan jumlah *hidden layer* pada *MLP* dapat disesuaikan dengan data yang digunakan, misalnya, berdasarkan jumlah data pelatihan. Sementara itu, *GMF* di dalam *Model NCF* bertugas untuk mengenali pola interaksi non-linier antara pengguna dan produk. Dengan demikian, penggabungan keduanya menghasilkan *model* yang lebih kuat dan efektif dalam rekomendasi rating produk kepada pengguna.

SIMPULAN

Kesimpulan yang diperoleh setelah melakukan serangkaian pengujian dengan beberapa skenario dan model berdasarkan data rating yang diambil dari MovieLens adalah sebagai berikut:

1. *Model NCF* mampu memberikan kinerja yang sangat memuaskan, terlihat dari tingginya nilai *recall* sebesar 69% dan *NDCG* sebesar 81%. Hal ini patut diperhatikan, terutama mengingat ketidakseimbangan data yang digunakan dalam eksperimen ini.
2. *Model NCF* ternyata sangat fleksibel dalam penggunaannya. Kemampuannya untuk menggunakan fitur-fitur tambahan seperti informasi pemeran, *genre*, atau pengarang film membuktikan bahwa *model* ini mampu menangkap pola-pola yang lebih kompleks dalam hubungan antara pengguna dan *film*.
3. Terbukti bahwa nilai parameter *L* pada metrik Top-*L* sangat berpengaruh terhadap hasil yang diperoleh, terutama dalam hal *recall*. Penggunaan nilai *L* yang umum seperti lima, sepuluh, atau bahkan dua puluh dalam pengujian menjadi pilihan yang efisien dan relevan.
4. Meskipun *Model NCF* memiliki kelemahan utama dalam menghadapi masalah pengguna baru yang memerlukan pelatihan ulang, namun hal ini dapat diatasi dengan cara yang cerdas. Penggabungan *Model NCF* membuka pintu solusi untuk masalah

ini, memungkinkan penggunaan model dalam skala yang jauh lebih besar dan lebih berkelanjutan.

DAFTAR PUSTAKA

- Alsini, A., Huynh, D. Q., dan Datta, A. (2020) Hit ratio: An Evaluation Metric for Hashtag Recommendation. *Jurnal Arxiv*, 1.
- Benjamin, W. (2021, Januari 18). *Ranking Evaluation Metrics for Recommender Systems*. Diakses 14 Juni 2023. Dari: <https://towardsdatascience.com/ranking-evaluation-metrics-for-recommender-systems-263d0a66ef54>.
- Brownlee, J. (2020, Agustus 26) *Train-Test Split for Evaluating Machine Learning Algorithms*. Diakses 14 Juni 2023. Dari: <https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>.
- Chen, D. (2020, Juli 8). *Recommender System — Matrix Factorization*. Diakses 13 Juni 2023. Dari: <https://towardsdatascience.com/recommendation-system-matrix-factorization-d61978660b4b>.
- Covington, P., Adams, J., dan Sargin, E. (2016) Deep Neural Networks for YouTube Recommendations. *The Proceedings of the 10th ACM Conference on Recommender Systems*, 191-198.
- Dhinakaran, A. (2023, Januari 26) *Demystifying NDCG*. Diakses 14 Juni 2023. Dari: <https://towardsdatascience.com/demystifying-ndcg-bee3be58cfe0>.
- Erlsha, E., dan Hiryanto, L. (2017) Sistem Rekomendasi Perencanaan Studi Mahasiswa dengan menggunakan Algoritma Apriori dan Naïve Bayes (Studi Kasus FTI UNTAR). *Jurnal JIKSI (Jurnal Ilmu Komputer dan Sistem Informasi)*, 3(1): 1.
- Fajriansyah, M., Adikara, P. P., dan Widodo, A. W., (2021) Sistem Rekomendasi Film Menggunakan Content Based Filtering. *Jurnal JPTIHK (Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer)*, 5(6): 2188-2199.
- Fitrianti, A. R., Rohmani, A., dan Widjanarto, W. (2020) Sistem Rekomendasi Film Berbasis Website Dengan Metode Prototype Menggunakan Algoritma K-Nearest Neighbors (KNN). *Jurnal JOINS*, 5(1): 8.
- Harper, F. M., dan Konstan, J. A. (2015) The MovieLens Datasets: History and Context. *Jurnal TiiS (Transactions on Interactive Intelligent Systems)*, 5(19): 1-10. DOI: 10.1145/2827872.
- He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T. (2020) Neural Collaborative Filtering. *Jurnal Arxiv*, 4.
- Koren, Y., Bell, R., dan Volinsky, C. (2009) Matrix Factorization Techniques for Recommender Systems. *Jurnal Computer*, 42(8): 30-37. DOI: 10.1109/MC.2009.263.
- Longo, C. (2018, November 23) *Evaluation Metrics for Recommender Systems*. Diakses 14 Juni 2023. Dari: <https://towardsdatascience.com/evaluation-metrics-for-recommender-systems-df56c6611093>.
- NVIDIA. (n.d.). *Recommendation System*. Diakses 13 Juni 2023. Dari: <https://www.nvidia.com/en-us/glossary/data-science/recommendation-system/>.
- Parwita, W. G. S., Swari, M., dan Welda. (2018) Perancangan Sistem Rekomendasi Dokumen Dengan Pendekatan Content-based Filtering. *Jurnal CESS (Journal of Computer Engineering System and Science)*, 3(1): 65-74.
- Priya, S. K., Manonmani, T., Dharshana, N., dan Ragaanasuya, K. (2022) Movie recommendation system with hybrid collaborative and content-based filtering using convolutional neural network. *Jurnal IJHS (International Journal of Health Sciences)*, 174(6): 5357-5372.

- Pradhan, R., Swami, A. C., Saxena, A., dan Rajpoot, V. (2021) A Study on Movie Recommendations using Collaborative Filtering. *Jurnal IOP (Institute Of Physics)*, 1119(1): 2. DOI: 10.1088/1757-899X/1119/1/012018.
- Rapaport, E. (2022, Juni 12) *MovieLens-1M Deep Dive — Part I*. Diakses 13 Juni 2023. Dari: <https://towardsdatascience.com/movie-lens-1m-deep-dive-part-i-8acfeda1ad4>.
- Sandrya, V., Wasino, dan Arisandi, D. (2022) Sistem Rekomendasi Film Menggunakan Metode Multiple Attribute Utility Theory. *Jurnal Computatio: Journal of Computer Science and Information Systems*, 6(1): 19-30.
- Schafer, B., dan Frankowski, D. (2007) Collaborative Filtering Recommender Systems. *Jurnal ResearchGate*, 1.
- Shahawy, S. E. (2020, Januari 17). *Understanding Matrix Factorization for recommender systems*. Diakses 13 Juni 2023. Dari: <https://towardsdatascience.com/understanding-matrix-factorization-for-recommender-systems-4d3c5e67f2c9>.
- Sugiyanto, S., dan Muhammad, M. (2022) Item Based Collaborative Filtering Based on Highest Item Similarity. *Jurnal IJAIR (Intenational Journal of Artificial Intelligence Research)*, 6(1): 3.
- Wang, Y. (2022, Februari 14). *Netflix's Recommendation Systems: Entertainment Made for You*. Diakses 12 Juni 2023. Dari: <https://illumin.usc.edu/netflixs-recommendation-systems-entertainment-made-for-you/>
- Wibowo, D., E. dan Munir, Rinaldi. (2013). Sistem Rekomendasi Jual Beli Barang dengan Memanfaatkan Metode Collaborative Filtering dan Basis Data Graf (Studi Kasus : Bukalapak.com). *School of Electrical*
- Engineering and Informatics, Institute Technology of Bandung.*