

RESEARCH ARTICLE

Pengembangan Sistem Rekomendasi Anime Berbasis *Deep Q-Network* (DQN)

Ricky Ariesta Fakhruddin, Izzatul Ummah* and Selly Meliana

Fakultas Informatika, Universitas Telkom, Bandung, 40257, Jawa Barat, Indonesia

* Corresponding author: izzatulummah@telkomuniversity.ac.id

Abstrak

Sistem rekomendasi berperan penting dalam membantu pengguna menemukan konten sesuai preferensi di tengah banyaknya informasi. Penelitian ini membandingkan dua pendekatan, yaitu *Collaborative Filtering* (CF) berbasis *Singular Value Decomposition* (SVD) dan *Reinforcement Learning* (RL) berbasis *Deep Q-Network* (DQN). Fokus utama penelitian adalah menilai efektivitas keduanya dalam memberikan rekomendasi anime yang relevan, baik untuk pengguna eksisting maupun pengguna baru (*cold-start*). Dataset penelitian diambil dari Kaggle, melalui tahap preprocessing berupa pembersihan data, normalisasi fitur, dan *encoding genre* dengan *one-hot*. Model CF dilatih menggunakan parameter hasil tuning, sedangkan model RL dibangun dalam lingkungan simulasi dengan fungsi *reward* berbobot yang menggabungkan rating pengguna, skor global anime, dan kesamaan preferensi genre. Evaluasi dilakukan menggunakan skenario Top- N Recommendation ($N = 1, 3, 5, 10, 15, 20$) dengan metrik *Precision@N*, *Recall@N*, dan *F1-Score@N*. Item relevan untuk pengguna eksisting ditentukan berdasarkan *reward* persentil ke-80, sementara untuk pengguna baru ditetapkan pada anime dengan skor global 19.0. Hasil menunjukkan RL dengan DQN unggul pada masalah *cold-start*, sedangkan CF lebih baik untuk pengguna dengan riwayat interaksi. Perbandingan ini menyoroti kelebihan dan keterbatasan masing-masing pendekatan, sekaligus memberi panduan dalam memilih strategi rekomendasi sesuai konteks pengguna.

Key words: *Deep q-network*, *Reinforcement Learning*, Colaborative Filtering, Sistem Rekomendasi Anime

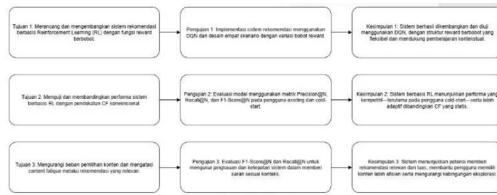
Pendahuluan

Sejak awal 1990-an, budaya otaku Jepang telah menyebar luas ke berbagai negara di Asia, seperti Hong Kong, Taiwan, dan daratan Tiongkok. Komik dan anime seperti *Slam Dunk* dan *Crayon Shin-Chan*, karakter seperti Hello Kitty, serta drama televisi seperti *Tokyo Love Story* berhasil menarik perhatian penggemar di luar Jepang. Saat ini, kebudayaan Jepang telah berkembang menjadi fenomena global. Berdasarkan data dari GoodStats, pada tahun 2023, Amerika Serikat menjadi negara dengan jumlah penonton anime terbesar kedua setelah Jepang, dengan proporsi lebih dari 70%. Hal ini mencerminkan bahwa animasi khas Jepang ini tidak hanya dikenal luas, tetapi juga memiliki basis penggemar yang signifikan di seluruh dunia. Namun, di tengah popularitas yang semakin meningkat, terdapat fenomena yang dikenal sebagai *content fatigue* atau *decision paralysis*, yaitu kesulitan pengguna dalam memilih konten akibat terlalu banyaknya pilihan yang tersedia. Dalam konteks menonton anime, banyaknya judul dari berbagai genre tersedia di berbagai *platform streaming*. Banyaknya pilihan ini sering kali membuat pengguna kewalahan dalam menentukan

tontonan, sehingga mereka cenderung menghabiskan waktu lebih banyak berusaha untuk memilih tontonan daripada menikmati konten itu sendiri.

Untuk mengatasi masalah ini, sistem rekomendasi merupakan salah satu solusi yang penting di era digital. Sistem rekomendasi pertama kali diperkenalkan pada pertengahan 1990-an dan kini telah diterapkan secara luas, contohnya pada *platform* perdagangan elektronik seperti Tokopedia dan Shopee, di mana pengguna diberikan rekomendasi produk yang dipersonalisasi berdasarkan preferensi dan riwayat pencarian mereka. Selain itu, minat penelitian di bidang ini terus meningkat karena potensi sistem rekomendasi untuk menyelesaikan masalah kelebihan informasi sekaligus membantu pengguna membuat keputusan yang lebih efektif. Salah satu pendekatan yang relevan dalam penelitian ini adalah penerapan sistem rekomendasi pada pemilihan tontonan anime. Dengan banyaknya pilihan yang tersedia, salah satunya, sistem rekomendasi berbasis *Collaborative Filtering* (CF) telah digunakan untuk menganalisis pola interaksi pengguna.

Namun, metode ini memiliki keterbatasan, terutama ketika data yang tersedia bersifat jarang (*sparse data*) atau ketika pengguna baru bergabung dengan *platform* (*cold-start problem*). Sebagai alternatif, pendekatan berbasis *Reinforcement Learning* (RL), khususnya dengan



GAMBAR 1

Gambar 1. Keterkaitan antara tujuan, pengujian dan kesimpulan.

algoritma *Deep QNetwork* (DQN), menawarkan solusi yang lebih adaptif. Pendekatan ini memungkinkan sistem rekomendasi untuk belajar secara dinamis dari interaksi pengguna dan menyesuaikan rekomendasi berdasarkan perubahan preferensi pengguna dari waktu ke waktu. Penelitian ini bertujuan untuk membandingkan performa metode CF berbasis *Singular Value Decomposition* (SVD) dengan metode RL berbasis DQN dalam konteks rekomendasi tontonan anime. Diharapkan, pendekatan RL ini dapat menghasilkan rekomendasi yang lebih relevan dan akurat, sekaligus mengatasi permasalahan *cold-start* dan *content fatigue*, sehingga dapat meningkatkan pengalaman pengguna dalam memilih anime yang sesuai dengan preferensinya.

Topik dan Batasannya

Topik penelitian ini berfokus pada pengembangan dan evaluasi sistem rekomendasi tontonan anime berbasis *Reinforcement Learning* (RL) dengan pendekatan *Deep QNetwork* (DQN), serta perbandingannya dengan metode *Collaborative Filtering* (CF). Penelitian ini dibatasi oleh beberapa hal sebagai berikut:

1. Dataset yang digunakan adalah dataset publik dari Kaggle, berisi data rating dan preferensi pengguna terhadap anime.
2. Metode CF yang digunakan dapat berupa pendekatan *memory-based* (analisis pola interaksi antar pengguna dan item) maupun *model-based* menggunakan *Singular Value Decomposition* (SVD).
3. Sistem rekomendasi utama menggunakan pendekatan RL berbasis DQN, tanpa integrasi eksplisit dengan CF.
4. Pengujian dilakukan dengan skenario top- N recommendation menggunakan nilai $N = 1, 3, 5, 15$, dan 20 .
5. Evaluasi performa menggunakan *Precision*, *Recall*, dan *F1-score* untuk menilai relevansi dan keseimbangan hasil rekomendasi.
6. Proses training pada DQN hanya menggunakan sampel dengan jumlah pengguna tertentu, disesuaikan dengan keterbatasan sumber daya komputasi yang tersedia.

Penelitian ini bertujuan merancang dan mengembangkan sistem rekomendasi tontonan anime berbasis *Reinforcement Learning* dengan algoritma DQN yang memanfaatkan fungsi *reward* berbobot agar mampu menghasilkan rekomendasi yang sesuai dengan preferensi pengguna. Selain itu, penelitian ini juga bertujuan membandingkan performa metode DQN dengan CF untuk menentukan teknik yang lebih optimal dalam konteks dataset yang digunakan. Lebih jauh, penelitian ini juga diarahkan untuk mengatasi permasalahan *content fatigue* pada teknologi dan hiburan, dengan memberikan solusi berupa sistem rekomendasi yang dapat membantu pengguna memilih tontonan anime yang relevan, menarik, dan sesuai dengan preferensi mereka.

Tinjauan Pustaka

Budaya anime telah menjadi fenomena global sejak tahun 1990-an. Sebagai bagian dari budaya populer Jepang, anime tidak hanya diterima di Jepang, tetapi juga di berbagai negara lain, termasuk Amerika Serikat dan Hong Kong [1]. Menurut data dari GoodStats (2023), anime

merupakan salah satu bentuk hiburan digital yang paling banyak diakses. Amerika Serikat mencatat lebih dari 70% penonton anime global, menjadikannya pasar terbesar setelah Jepang. Fenomena ini juga terlihat di Indonesia, di mana budaya anime diadaptasi ke dalam bentuk ekspresi lokal seperti cosplay. Studi oleh Putri (2016) di Bandung menemukan bahwa komunitas cosplay berkembang sebagai sarana ekspresi budaya populer yang memadukan elemen lokal dengan karakter anime [2]. Hal ini menegaskan bahwa anime tidak hanya sekadar hiburan, tetapi juga membentuk identitas sosial dan komunitas global yang erat. Sistem rekomendasi mulai berkembang sejak pertengahan 1990-an dan kini menjadi elemen penting dalam ekosistem digital [3]. Sistem ini berfungsi membantu pengguna memilih konten yang relevan dengan preferensi mereka. Dua pendekatan utama yang umum digunakan adalah *Content-Based Filtering* dan *Collaborative Filtering* (CF).

Collaborative Filtering

CF tidak bergantung pada fitur item, melainkan pada pola interaksi pengguna-item. Prinsip utamanya adalah bahwa pengguna dengan preferensi serupa di masa lalu kemungkinan menyukai item yang sama di masa depan. CF terbagi menjadi:

1. **User-Based CF:** Pada pendekatan userbased, prediksi rating yang diberikan oleh pengguna u terhadap item i dihitung menggunakan rata-rata tertimbang dari rating yang diberikan oleh pengguna lain v , yang dianggap mirip dengan u :

$$\hat{r}_{u,i} = \bar{r}_u + \frac{\sum_{v \in N(u)} \text{sim}(u, v) \cdot (r_{v,i} - \bar{r}_v)}{\sum_{v \in N(u)} |\text{sim}(u, v)|} \quad (1)$$

2. **Item-Based CF:** Pada pendekatan itembased, prediksi rating diberikan dengan menghitung kemiripan antara item i dan item lain j yang telah dirating oleh pengguna u .

$$\hat{r}_{u,i} = \frac{\sum_{j \in S(i)} \text{sim}(i, j) \cdot r_{u,j}}{\sum_{j \in S(i)} |\text{sim}(i, j)|} \quad (2)$$

Kelebihan CF adalah kemampuannya menemukan rekomendasi yang lebih beragam, namun rentan terhadap masalah data jarang (*sparsity*). Untuk mengatasi hal ini, pendekatan model-based seperti *Matrix Factorization* (MF) digunakan. Salah satu algoritma populer adalah *Singular Value Decomposition* (SVD) [4], yang memfaktorkan matriks rating menjadi dimensi laten pengguna dan item.

SVD mampu menghasilkan prediksi rating akurat meski data jarang, serta stabil terhadap overfitting melalui regularisasi. Proses ini mencakup: *preprocessing dataset*, pembentukan matriks *sparse*, pelatihan dengan algoritma SVD, prediksi rating, dan evaluasi menggunakan metrik RMSE, MAE, *Precision@N*, *Recall@N*, serta *F1-Score*.

Reinforcement Learning dan Deep Q-Network (DQN)

Reinforcement Learning (RL) memungkinkan agen belajar dari interaksi dengan lingkungan melalui mekanisme trial-and-error, dengan *reward* sebagai umpan balik [5][6]. Dalam konteks sistem rekomendasi, RL dimodelkan sebagai Markov Decision Process (MDP) dengan:

1. **State (s):** representasi preferensi dinamis pengguna,
2. **Action (a):** item (anime) yang direkomendasikan,
3. **Reward (r):** feedback dari pengguna,
4. **Transition ($s \rightarrow s'$):** perubahan preferensi setelah aksi.

Dasar dari RL adalah Persamaan Bellman [7], yang memperbarui nilai fungsi Q secara iteratif dengan menggabungkan *reward* saat ini dan perkiraan *reward* masa depan. Algoritma Q -Learning dan *Deep Q-Network* (DQN) dibangun di atas konsep ini, di mana nilai Q dipelajari melalui jaringan saraf dalam.

Table 1. labelformat=empty

Keterkaitan Antara Tujuan, Pengujian dan Kesimpulan			
Nomor	Tujuan Penelitian	Pengujian	Kesimpulan
1	Merancang dan mengembangkan sistem rekomendasi berbasis <i>Reinforcement Learning</i> (RL) dengan fungsi <i>reward</i> berbobot.	Implementasi sistem rekomendasi menggunakan DQN dan desain empat skenario dengan variasi bobot <i>reward</i> .	Sistem berhasil dikembangkan dan diuji menggunakan pendekatan Deep QNetwork, dengan struktur <i>reward</i> berbobot yang fleksibel dan mendukung pembelajaran kontekstual.
2	Menguji dan membandingkan performa sistem berbasis RL dengan pendekatan CF konvensional.	Evaluasi model menggunakan metrik, $Precision@N$, $Recall@N$, dan $F1Score@N$ pada pengguna existing dan <i>cold-start</i> .	Sistem berbasis RL menunjukkan performa yang kompetitif, khususnya pada pengguna <i>cold-start</i> , serta lebih adaptif terhadap variasi pengguna dibandingkan CF yang statis.
3	Mengurangi beban pemilihan konten dan mengatasi content fatigue melalui rekomendasi yang relevan.	Evaluasi $F1Score@N$ dan $Recall@N$ untuk mengukur jangkauan dan ketepatan sistem dalam memberikan saran yang sesuai konteks.	Sistem menunjukkan potensi dalam memberikan rekomendasi yang relevan dan luas, membantu pengguna memilih konten dengan lebih efisien serta mengurangi kebingungan dalam eksplorasi.

Untuk menentukan nilai optimal dari suatu aksi dalam kondisi tertentu, digunakan persamaan Bellman Equation [7], yang menjadi dasar bagi algoritma RL seperti *Q* Learning dan DQN. Persamaan Bellman secara umum dituliskan sebagai:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[r + \gamma \max_{a'} Q(s', a') - Q(s, a) \right] \quad (3)$$

DQN menggunakan parameter θ (bobot dan bias jaringan) untuk memprediksi nilai Q , yang kemudian dibandingkan dengan target Bellman. Selisihnya diminimalkan menggunakan fungsi Mean Squared Error (MSE) dan optimisasi algoritma Adam. Strategi epsilongreedy digunakan agar agen seimbang antara eksploitasi (memilih item dengan Q -value tertinggi) dan eksplorasi (memilih item acak). Kelebihan DQN dalam sistem rekomendasi adalah:

1. Adaptif: mampu menyesuaikan rekomendasi terhadap preferensi pengguna yang berubah,
2. Jangka Panjang: mempertimbangkan *reward* masa depan, bukan hanya feedback instan,
3. Fleksibel: dapat menggabungkan sinyal rating pengguna (CF), skor popularitas global, dan kesamaan konten dalam fungsi *reward* berbobot [5][17][18].

Pendekatan Reinforcement Learning-Deep QNetwork Dalam Penelitian Ini

Dalam penelitian ini, sistem rekomendasi dikembangkan murni menggunakan RL berbasis DQN tanpa integrasi eksplisit dengan CF. Komponen utamanya:

1. State: representasi pengguna dan preferensi dinamis berbasis vektor genre,
2. Action: kandidat anime dalam bentuk *one-hot* vector,
3. Reward: kombinasi tiga sinyal (rating user, skor global, dan kesamaan genre) dengan bobot α yang divariasikan,
4. Training: pembelajaran berbasis persamaan Bellman dengan update Q -value,
5. Evaluasi: $Precision@N$, $Recall@N$, dan $F1Score@N$ pada pengguna existing serta $Precision@N$ pada *cold-start* user.

Dengan rancangan ini, penelitian berfokus pada pembuktian bahwa DQN dapat menghasilkan rekomendasi yang relevan, adaptif, dan mampu mengurangi fenomena *content fatigue* dalam konsumsi konten anime.

Metodologi Penelitian

Dalam perancangan dan pengaplikasian sistem rekomendasi anime, proses pembangunan dibagi ke dalam beberapa tahapan. Secara garis besar, sistem dibangun dalam dua pendekatan: *Collaborative Filtering* (CF) berbasis algoritma *Singular Value Decomposition* (SVD), dan *Reinforcement Learning* (RL) berbasis *Deep Q-Network* (DQN). Tahapan pembangunan meliputi pendahuluan, pengumpulan data, preprocessing, perancangan model, verifikasi dan validasi, serta desain perangkat keras dan lunak.

Tahap Pendahuluan

Tahap awal penelitian ini bertujuan mengatasi keterbatasan sistem rekomendasi konvensional, khususnya dalam hal:

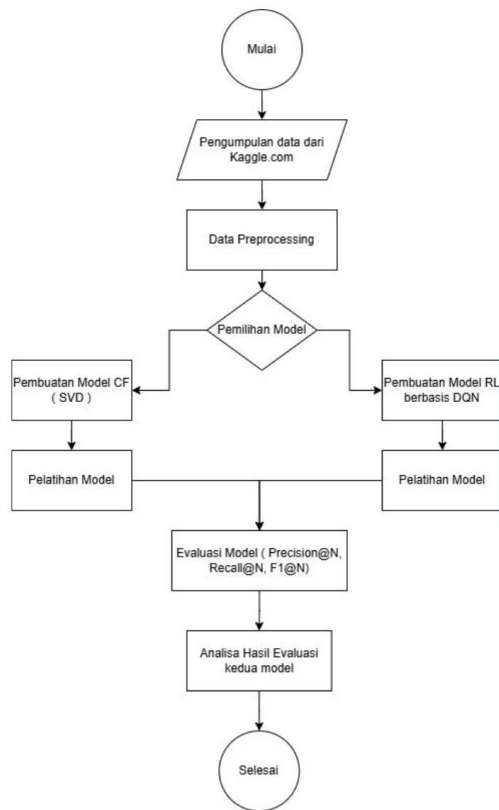
1. Peningkatan akurasi prediksi rating,
2. Relevansi rekomendasi yang diberikan,
3. Adaptasi terhadap pengguna baru (*cold-start problem*).

Fokus utama penelitian adalah membangun dan membandingkan dua pendekatan sistem rekomendasi:

1. CF berbasis SVD (CF-SVD) sebagai baseline,
2. RL berbasis DQN (RL-DQN) dengan fungsi *reward* berbobot.

Kedua pendekatan dikembangkan secara terpisah, tanpa digabungkan dalam arsitektur hybrid, agar perbedaan performa dapat dievaluasi secara langsung. Pengembangan dilakukan di Google Colab menggunakan Python, dengan library:

1. NumPy, Pandas → manipulasi data,
2. Scikit-learn → *preprocessing* dan evaluasi,
3. Surprise → implementasi CF-SVD,
4. PyTorch → pembangunan DQN,



Gambar 2. Proses Perancangan Sistem Rekomendasi

5. *pickle, gc, itertools, random, collections* → mendukung manajemen data dan optimisasi model.

Tahap Perancangan Sistem

Dataset diperoleh dari Kaggle dalam bentuk data interaksi pengguna terhadap anime (*user_id*, *anime_id*, *user_rating*). Dataset ini kemudian diperkaya dengan metadata anime (judul, genre, skor, jumlah episode, tipe anime). Data ini menjadi dasar untuk analisis pola preferensi dan pelatihan model rekomendasi yang dapat dilihat pada gambar 2.

Tahap Preprocessing Data

Tahapan preprocessing meliputi:

1. Penggabungan dataset rating pengguna dengan metadata anime melalui *anime_id*,
2. Penanganan :

missing value

- Kolom episodes diubah menjadi numerik, nilai hilang diganti median,
 - Kolom *type* dan genre kosong diberi label "Unknown",
 - Kolom *anime_score* kosong diganti dengan rata-rata skor.
3. Transformasi fitur genre dengan *one-hot encoding* menggunakan *MultiLabelBinarizer* → menghasilkan matriks biner efisien (int8).
 4. Penghapusan kolom genre asli untuk mencegah redundansi.

Output preprocessing menghasilkan dataset bersih dan terstruktur, yang digunakan sebagai input untuk dua model:

1. CF-SVD → menggunakan *user_id*, *anime_id*, *user_rating*,

2. RL-DQN → menggunakan *user_rating*, *anime_score*, dan vektor genre.

Tahap Verifikasi dan Validasi

Verifikasi → memastikan implementasi algoritma sesuai rancangan:

1. CF-SVD: proses training & prediksi,
2. RL-DQN: alur *state-action-reward*, perhitungan Qvalue, dan sampling rekomendasi. Validasi → memastikan sistem memberikan hasil akurat dan relevan:
3. *Ground truth* ditentukan berdasarkan persentil ke-80 dari distribusi rating pengguna → hanya item dengan rating $\geq \text{threshold}$ dianggap relevan.
4. Metrik evaluasi:

- RMSE & MAE → untuk *error* prediksi numerik (CF),
- *Precision@N*, *Recall@N*, *F1@N* → untuk relevansi rekomendasi (RL & CF).

Desain Perangkat Keras dan Lunak

Google Colab digunakan sebagai platform dengan dasar:

1. Lingkungan *cloud* yang fleksibel (integrasi *Google Drive*),
2. Dukungan GPU NVIDIA L4 untuk training RL,
3. Efisiensi waktu & memori.

Untuk RL, jumlah pengguna dalam training dibatasi ± 2000 pengguna acak untuk menjaga waktu pelatihan tetap realistis dan menghindari kehabisan memori.

Untuk penentuan *ground truth* dilakukan dengan mencari item relevan ditentukan dengan algoritma:

```

Algoritma 1: Penentuan relevansi item
assert 'user_rating' in dataset_full.columns, "Missing 'user_rating' column"
percentile_80 = dataset_full['user_rating'].quantile(0.80)
rating_threshold = percentile_80
  
```

Hanya item dengan rating $\geq$ persentil ke-80 yang dianggap relevan → sekitar 20% anime dengan rating tertinggi.

Arsitektur sistem rekomendasi dibangun dengan:

1. *Collaborative Filtering* (SVD)
 - *Matrix Factorization* berbasis SVD,
 - Menangkap representasi laten pengguna & item,
 - Dievaluasi dengan RMSE, MAE, *Precision@N*, *Recall@N*, *F1-Score@N*.
2. *Reinforcement Learning - Deep Q-Network* (DQN)
 - Model rekomendasi dimodelkan sebagai Markov Decision Process (MDP),
 - Q-Network mempelajari nilai pasangan (state, action),
 - Training dengan epsilon-greedy & optimisasi Adam,
 - *Reward* = kombinasi berbobot dari *user_rating*, *anime_score*, *genre.similarity*.

Hasil dan Pembahasan

Evaluasi

Evaluasi dilakukan untuk menilai performa sistem rekomendasi berbasis *Collaborative Filtering* (CF-SVD) dan *Reinforcement Learning* (RL-DQN). Dua jenis pengguna menjadi fokus utama:

Table 2. labelformat=empty

Skenario	Batch	Epoch	Avg Precision	Avg Recall	Avg F1score
Skenario 1	128	5	0.3768	0.5567	0.3342
Skenario 2	128	5	0.3675	0.5618	0.3319
Skenario 3	64	5	0.4629	0.6174	0.3959
Skenario 4	64	10	0.3166	0.5149	0.2915

1. Pengguna existing, adalah pengguna yang sudah memiliki riwayat interaksi, sehingga evaluasi menggunakan *Precision@N*, *Recall@N*, *F1-Score@N*, rmse, maE.
2. Pengguna *cold-start*, adalah pengguna yang tidak memiliki riwayat interaksi, sehingga evaluasi menggunakan *Precision@N* saja. Metrik dihitung pada berbagai nilai $Top - N(1, 3, 5, 10, 15, 20)$ untuk menilai performa sistem dari rekomendasi. Ini

Analisis pada sistem rekomendasi RL-DQN dilakukan berdasarkan perbandingan metrik antar konfigurasi pelatihan, yaitu:

1. Jumlah epoch: 5 dan 10 .
2. Ukuran batch: 64, 128 , dan 256 .

Tujuannya adalah untuk melihat pengaruh perubahan fungsi *reward* terhadap:

1. Kinerja sistem dalam merekomendasikan anime yang relevan.
2. Efektivitas personalisasi pada pengguna aktif.
3. Kemampuan mengatasi masalah *cold-start*.

Hasil Evaluasi dan Analisis

Hasil eksperimen memperlihatkan perbedaan antar pendekatan.

1. Penekanan pada skor global (Skenario 3) terbukti paling efektif karena mampu menyeimbangkan personalisasi dan kualitas kolektif.
2. Pengguna *cold-start*: CF dapat menghasilkan rekomendasi akurat dengan memanfaatkan skor global, namun tidak personal. RL-DQN dengan bobot skor global (Skenario 3) memberikan hasil unggul (*Precision@1* mencapai 0.93), sedangkan RL berbasis rating personal (Skenario 4) gagal menghadapi *cold-start*.
3. Pengaruh parameter pelatihan: Penambahan epoch umumnya meningkatkan Recall dan F1, tetapi beresiko overfitting pada *cold-start*. Batch size 128 memberi keseimbangan terbaik, sementara batch besar (256) kadang meningkatkan recall, namun menurunkan precision.

Secara umum, eksperimen membuktikan bahwa pendekatan RL-DQN dengan penekanan skor global adalah konfigurasi paling efektif, baik untuk pengguna aktif maupun *cold-start*. Sementara itu, CF tetap relevan sebagai baseline sederhana, dan konfigurasi berbasis rating personal memiliki keterbatasan untuk konteks pengguna.

Kesimpulan

Penelitian ini berhasil merancang dan mengimplementasikan sistem rekomendasi berbasis *Reinforcement Learning* (RL) menggunakan algoritma Deep *Q*-Network (DQN). Sistem dirancang untuk memanfaatkan fungsi *reward* berbobot yang terdiri dari tiga komponen utama: rating pengguna (user rating), skor global anime (anime score), dan kesamaan preferensi genre (genre similarity). Pendekatan ini bertujuan menghasilkan rekomendasi yang adaptif terhadap preferensi individual pengguna serta lebih efektif dalam menangani permasalahan

cold-start, dibanding pendekatan konvensional seperti *Collaborative Filtering* (CF). Meskipun penelitian ini mengacu pada inspirasi dari pendekatan kolaboratif, sistem yang dikembangkan tidak menggabungkan prediksi CF secara eksplisit, melainkan belajar langsung dari data interaksi pengguna melalui *reward* function yang dirancang khusus. Dengan demikian, proses pembelajaran RL berjalan secara independen, memungkinkan sistem untuk membentuk strategi rekomendasi yang lebih personal, fleksibel, dan tidak hanya terbatas pada pola historis.

Berdasarkan proses implementasi dan evaluasi yang telah dilakukan, seluruh rumusan masalah dalam penelitian ini telah terjawab. Sistem dibangun menggunakan model DQN, dan CF berbasis Singular Value Decomposition (SVD) digunakan secara terpisah sebagai baseline pembandingan. Evaluasi dilakukan terhadap dua kelompok pengguna, yaitu pengguna eksisting dan pengguna baru (*cold-start*), digunakan metrik - metrik seperti, *Precision@N*, *Recall@N*, dan *F1-Score @ N* dengan rentang nilai N dari 1 hingga 20. Untuk pengguna eksisting, item relevan dalam data uji ditentukan secara dinamis berdasarkan fungsi *reward* system itu sendiri, yaitu dengan menganggap item sebagai relevan apabila nilai *reward*-nya berada pada atau di atas persentil ke80 dari distribusi *reward* hasil evaluasi terhadap data uji. Pendekatan ini memastikan bahwa proses evaluasi tetap konsisten dan kontekstual dengan perilaku model dan strategi pembelajaran yang digunakan.

Ambang batas relevansi item dalam evaluasi sistem ditentukan berdasarkan analisis distribusi rating pada dataset. Nilai threshold relevansi diambil dari persentil ke-80 pada kolom *user_rating*, yang menghasilkan nilai ≥ 9.0 . Artinya, hanya 20% rating tertinggi dari seluruh data yang dianggap mewakili item dengan relevansi tinggi. Pendekatan ini diterapkan secara konsisten pada semua jenis pengguna, baik pengguna dengan riwayat interaksi maupun pengguna baru, sehingga proses evaluasi selaras dengan kriteria kualitas item berdasarkan penilaian komunitas secara agregat. Hasil evaluasi menunjukkan bahwa masing-masing skenario dengan variasi bobot *reward* memberikan performa yang berbeda, baik pada pengguna eksisting maupun pengguna *cold-start*. Hal ini menunjukkan bahwa komponen *reward* yakni bobot rating pengguna (a_1), bobot skor global (a_2), dan bobot kesamaan genre (a_3) berkontribusi secara berbeda terhadap arah pembelajaran dan hasil rekomendasi.

Perbedaan yang paling menonjol terlihat pada skenario yang menekankan bobot skor global (a_2), yang menghasilkan peningkatan signifikan terutama pada metrik *Recall@N* dan *F1-Score@N* untuk pengguna eksisting, serta pada *Precision@N* untuk pengguna *cold-start*. Hal ini menunjukkan bahwa preferensi pengguna dalam sistem tidak semata-mata bergantung pada item yang pernah dinilai, tetapi juga dipengaruhi oleh informasi eksternal seperti skor global dan kesamaan genre yang dipertimbangkan dalam *reward*. Pada pengguna baru, perubahan bobot *reward* memberikan dampak yang lebih signifikan terhadap hasil evaluasi. Karena tidak memiliki riwayat interaksi, sistem sepenuhnya mengandalkan sinyal awal dari fungsi *reward*, terutama skor global dan kesamaan genre. Dalam konteks ini, variasi bobot a secara langsung memengaruhi arah pembelajaran dan kualitas rekomendasi.

Hasil evaluasi menunjukkan bahwa model dengan penekanan pada skor global (a_2) atau kesamaan genre (a_3) mampu menghasilkan rekomendasi yang lebih relevan untuk pengguna baru, sedangkan dominasi bobot rating pengguna (a_1) justru menghambat performa, karena kurang memberikan informasi yang berguna dalam situasi *cold-start*. Secara keseluruhan, sistem rekomendasi yang dikembangkan dengan pendekatan *Reinforcement Learning* berbasis DQN menunjukkan kemampuan adaptif terhadap berbagai kondisi pengguna. Fungsi *reward* berbobot menjadi komponen kunci dalam mengatur keseimbangan antara personalisasi, popularitas, dan kesesuaian konten. Namun, performa sistem terbukti sangat bergantung pada konfigurasi bobot

reward yang digunakan. Temuan ini membuka peluang untuk pengembangan fungsi *reward* yang lebih kontekstual dan adaptif, guna meningkatkan akurasi rekomendasi baik pada pengguna eksisting maupun *cold-start* secara berkelanjutan.

Daftar Pustaka

1. Hatami W. Popular culture of Japanese anime in the digital age and the impact on nationalism of young Indonesian citizens. *Journal of Social Studies (JSS)*. 2018;14(1):1-10.
2. Putri A. Cosplay sebagai identitas budaya populer; 2016. Repository Universitas Telkom.
3. Ricci F, Rokach L, Shapira B. *Recommender Systems Handbook*. 2nd ed. New York, NY: Springer; 2015.
4. Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. *Computer*. 2009.
5. Lin Y, Liu Y, Lin F, Zou L, Wu P, Zeng W, et al. A survey on reinforcement learning for recommender systems. *arXiv preprint arXiv:210910665*. 2021.
6. Afsar MM, Crump T, Far BH. Reinforcement learning based recommender systems: A survey. *arXiv preprint arXiv:210106286*. 2021.
7. Roderick M, MacGlashan J, Tellex S. Implementing the Deep Q-Network. *arXiv preprint arXiv:171107478*. 2017.
8. Chen H, Dai X, Cai H, Zhang W, Wang X, Tang R, et al. Large-scale interactive recommendation with tree-structured policy gradient. *arXiv preprint arXiv:181105869*. 2018.
9. Chen M, Xu C, Gatto V, Jain D, Kumar A, Chi E. Off-policy actor-critic for recommender systems. *Proceedings of ACM*. 2019.
10. Gao C, Zheng Y, Li N, Li Y, Qin Y, Piao J, et al. A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Computing Surveys*. 2021.
11. Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv preprint arXiv:180101290*. 2018.
12. Kim S, Jeon S, Lee J, Kang J. Diversely regularized matrix factorization for accurate and aggregately diversified recommendation. *arXiv preprint arXiv:221101328*. 2022.
13. Lei Y, Li W. When collaborative filtering meets reinforcement learning. *arXiv preprint arXiv:190200715*. 2019.
14. Li C, Xia L, Ren X, Ye Y, Xu Y, Huang C. Graph transformer for recommendation. *arXiv preprint arXiv:230602330*. 2023.
15. Liang K, Zhang G, Guo J, Li W. An actor-critic hierarchical reinforcement learning model for course recommendation. *Mathematics*. 2022;10(18):3313.
16. Lv Z, Tong X. A reinforcement learning list recommendation model fused with graph neural networks. *Electronics*. 2023;12(18):3748.
17. Nadhifah AR, Wibowo AT. Sistem pemberi rekomendasi anime menggunakan pendekatan hybrid;. Fakultas Informatika, Universitas Telkom.
18. Nurfauzi AI, Wibowo AT. Sistem rekomendasi skincare menggunakan matrix factorization dengan metode non-negative matrix factorization;. Fakultas Informatika, Universitas Telkom.
19. Patoulia AA, Kiourtis A, Mavrogiorgou A, Kyriazis D. A comparative study of collaborative filtering in product recommendation. *Emerging Science Journal*. 2022;7(1):1-15.
20. Rahmawati S, Nurjanah D, Rismala R. Analisis dan implementasi pendekatan hybrid untuk sistem rekomendasi pekerjaan dengan metode knowledge-based dan collaborative filtering. *Indonesian Journal on Computing (Indo-JC)*. 2018;3(2):11-20.
21. Roziqin NM, Faisal M. Sistem rekomendasi pemilihan anime menggunakan user-based collaborative filtering. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*. 2024;9(1):1-10.
22. Zhang W, Wang J, Yu Y. Serendipitous personalized ranking for top-N recommendation. In: *Proceedings of RecSys*; 2012. p. 23-30.
23. Zhao X, Zhang Y, Xia L, Wu J. Deep reinforcement learning for list-wise recommendations. In: *Proceedings of the 25th ACM SIGKDD Conference*; 2019. p. 2769-77.
24. Zheng G, Noroozi V, Yu PS. Recommendations with negative feedback via pairwise deep reinforcement learning. In: *Proceedings of WSDM*; 2021. p. 620-8.
25. Zivic P, Vazquez H, Sanchez J. Scaling sequential recommendation models with transformers. *arXiv preprint*. 2021.
26. Zou L, Zhang P, Zhang D. Reinforcement learning to optimize long-term user engagement in recommender systems. *arXiv preprint arXiv:190205570*. 2019.