

Rekomendasi Video Game Menggunakan Metode Collaborative Filtering dengan K-NN

Kendrick Raphael Ticoalu^{a1}, Ida Ayu Gde Suwiprabayanti Putra^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹ticoalu.2208561097@student.unud.co.id
²iagsuwiprabayantiputra@unud.ac.id (Corresponding Author)

Abstract

As the digital age progresses, the more technology affects various aspects in our lives for example entertainment through video games. A problem arises where there are too many video games to choose from, so there is a need to find methods to narrow down the choices. This study implements a collaborative filtering-based video game recommendation system to analyze user preferences based on playtime data. The system processes user-game interaction data from a secondary dataset containing 14.3 million players and 50.9 million games, constructing a sparse matrix to map user playtime behavior. By identifying similar users through kNN, the system recommends games frequently played by users with comparable preferences. Evaluation on 100 sample users achieved an impressive mean precision of 88.12%, indicating that most recommended games were among the users' top 20 most-played titles. This study hopes to further enable people in finding more fun experiences in their lives.

Keywords: User-Based Collaborative Filtering, Recommendation System, Video Games, K-Nearest Neighbor

1. Pendahuluan

Pada masa kini, teknologi digital telah berkembang sangat pesat bahkan sampai mempengaruhi berbagai aspek kehidupan. Salah satu hal dimana teknologi mempengaruhi kehidupan kita sehari-hari ialah sebagai sumber hiburan melalui video game. Video game menjadi sumber hiburan utama untuk masyarakat luas sehingga industri video game mencapai 174.9 miliar dolar AS pada tahun 2020 menurut laporan *statista.com* [1]. Tetapi pertumbuhan tersebut memunculkan masalah dimana para konsumen terdapat sangat banyak pilihan video game.

Recommendation system merupakan suatu sistem yang digunakan untuk membantu proses penentuan suatu objek atau bahan penelitian lainnya [2]. Sistem rekomendasi menghasilkan informasi kumpulan data yang mungkin bermanfaat untuk penggunaan berdasarkan preferensi. Sistem rekomendasi bisa diterapkan untuk mengecilkan pilihan video game untuk dimainkan.

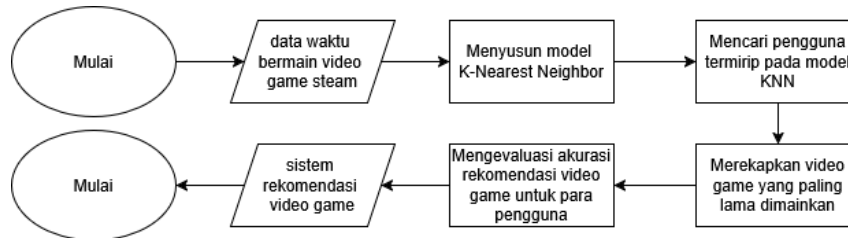
Penerapan sistem rekomendasi untuk video game sudah pernah diimplementasikan oleh Pramesti pada tahun 2022. Dalam penelitian tersebut, pendekatan Content-Based Filtering berdasarkan genre diterapkan sehingga mencapai presisi mean 87.75% [3]. Penelitian tersebut akan digunakan sebagai landasan dimana untuk menyusun sebuah sistem rekomendasi *video game* dengan pendekatan *Collaborative Filtering*.

Collaborative filtering adalah teknik yang digunakan untuk melakukan prediksi dengan memanfaatkan informasi tentang item yang disukai atau dipilih oleh pengguna lain [4]. Dalam penelitian ini *Collaborative Filtering* digunakan untuk menyusun sistem rekomendasi video game berdasarkan pengguna dengan waktu jam bermain (*playtime*) *video game* yang mirip. Teknik *Collaborative Filtering* mengimplementasi *K-Nearest Neighbor* untuk menganalisis preferensi pengguna yang mirip berdasarkan penempatannya dalam *feature space*.

2. Metode Penelitian

2.1. Desain Penelitian

Gambaran umum alur penelitian bisa direpresentasikan dalam flowchart berikut.



Gambar 1. Flowchart Alur Sistem

2.2. Data

Data adalah data sekunder yaitu data video game, pemain, dan rekapan lama waktu pemain bermain video game yang berasal dari Kaggle pada tautan berikut:

<https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset/data>

Data terdiri dari informasi 14.3 juta pemain, 50.9 juta *video game*, dan 41.2 juta ulasan pengguna. Data ulasan mengandung informasi apabila pengguna berkesan puas dengan *video game* dan waktu bermain sebagai berikut.

Tabel 1. Sampel Data Review

app_id	helpful	funny	date	is_recom mended	hours	user_id	review_i d
975370	0	0	2022- 12-12	TRUE	36.29 9999	51580	0
304390	4	0	2017- 02-17	FALSE	11.50 0000	2586	1
1085660	2	0	2019- 11-17	TRUE	336.5 0000 0	253880	2

Langkah *preprocessing* agar data dapat diolah dalam format yang bisa diterima sistem dilakukan dengan pembentukan matrix sparse. Matriks sparse dibuat dari data interaksi pengguna-item (dalam hal ini, jumlah jam bermain), di mana baris mewakili pengguna dan kolom mewakili item. Penyusunan matriks sparse diimplementasikan sebagai berikut.

```

user_ids = recommendations_df['user_id'].astype('category').cat.codes
item_ids = recommendations_df['app_id'].astype('category').cat.codes

unique_user_ids = recommendations_df['user_id'].astype('category').cat.categories
unique_item_ids = recommendations_df['app_id'].astype('category').cat.categories

user_game_matrix = coo_matrix((recommendations_df['hours'], (user_ids, item_ids)))
    
```

Maka melalui *preprocessing*, bisa dibentuk sebuah struktur data yang bertujuan untuk pemeriksaan *playtime* sebuah game untuk seorang pemain yang lebih efisien.

2.3. Collaborative Filtering

Collaborative Filtering menganalisis data matriks sparse pemain-video game yang berisi nilai lama jam bermain untuk ditempatkan pada sebuah *feature space*. Setiap axis adalah sebuah *video game* dimana setiap data poin pemain ditempatkan berdasarkan *playtime*. Dua pemain memiliki kemiripan apabila berada dalam lokasi yang dekat pada *feature space*.

Algoritma K-Nearest Neighbor yang melakukan proses klasifikasi dengan penggunaan mean data dari masing-masing kelas, kemudian menghitung jarak terdekat antara data uji dan mean data setia kelas tersebut [5]. K-NN diterapkan Collaborative Filtering untuk menentukan pemain-pemain yang mirip berdasarkan penempatannya pada *feature space*. Nilai *k* adalah jumlah *neighbor* yang ingin dicarikan yaitu pemain yang memiliki preferensi *video game* yang mirip. Nilai kemiripan pemain dihitung dengan berikut.

$$\text{cosine_similarity}(u, v) = (u \cdot v) / (||u|| * ||v||) \quad (1)$$

u : letak pemain pertama pada *feature space*
v : letak pemain kedua pada *feature space*

Collaborative Filtering untuk mencari pengguna yang mirip dengan preferensi yang mirip dengan pengguna terinput diimplementasikan sebagai berikut.

```
model_knn = NearestNeighbors(metric='cosine', algorithm='brute')
model_knn.fit(user_game_matrix)

distances, indices = model_knn.kneighbors(user_game_matrix.getrow(0), n_neighbors=6)
recommended_users = [unique_user_ids[i] for i in indices.flatten()[1:]]
```

2.4. Rekomendasi Game

Sistem rekomendasi yang dibangkitkan merekomendasikan *video game* berdasarkan asumsi bahwa *video game* yang paling menarik untuk seorang pengguna adalah *video game* yang paling banyak dimainkan oleh pengguna yang mirip. Maka sistem akan melakukan perhitungan total jam bermain yang dicapai pada setiap game oleh setiap pengguna mirip dan merekomendasikan *video game* dengan nilai total terbanyak. Pemilihan game diimplementasikan sebagai berikut.

```
def get_recommended_games_from_users(user_ids, recommendations_df, games_df, top_n=5):
    game_playtime = {}
    for user_id in user_ids:
        user_recs = recommendations_df[recommendations_df['user_id'] == user_id]
        for index, row in user_recs.iterrows():
            game_id = row['app_id']
            hours_played = row['hours']
            if game_id not in game_playtime:
                game_playtime[game_id] = 0
            game_playtime[game_id] += hours_played
    sorted_games = sorted(game_playtime.items(), key=lambda item: item[1], reverse=True)
    top_game_ids = [game_id for game_id, playtime in sorted_games[:top_n]]
    recommended_game_names = games_df[games_df['app_id'].isin(top_game_ids)]['title'].tolist()
    return recommended_game_names
```

Sistem menampilkan hasil analisis untuk merekomendasikan video game sebagai berikut:

```
Recommended games based on similar users ([np.int32(11422900), np.int32(7483536), np.int32(9284221), np.int32(6590596), np.int32(11161013)]):  
- Idle Champions of the Forgotten Realms  
- Insurgency  
- BorderLands 2  
- Far Cry 3  
- Cookie Clicker
```

Gambar 2. Output Rekomendasi Game

3. Hasil dan Diskusi

3.1. Data Testing

Sistem rekomendasi akan diuji akurasi dengan cara menginput sampel pengguna hanya dengan 10 *video game* yang paling lama dimainkan. Akurasi bertambah jika sistem rekomendasi mampu memilih 5 *video game* yang sesungguhnya termasuk *video game* yang paling dimainkan pengguna itu.

3.2. Metrik Evaluasi

Evaluasi menggunakan metrik evaluasi presisi untuk mengukur akurasi sistem rekomendasi. Dimana setiap *game* yang direkomendasi dimana sesungguhnya salah *video game* favorit yaitu yang paling sering dimainkan, sementara pemilihan *game* yang tidak termasuk ranking tersebut akan menurunkan nilai.

Maka presisi dikalkulasi dengan persamaan berikut.

$$precision = TP / (TP + FP) \quad (2)$$

TP : jumlah rekomendasi *video game* favorit pengguna
FP : jumlah rekomendasi *video game* yang bukan favorit pengguna

Perhitungan presisi diimplementasikan sebagai berikut.

```
def calculate_precision(user_id, recommended_games, recommendations_df, top_k_precision):  
  
    user_recs = recommendations_df[recommendations_df['user_id'] == user_id]  
    user_top_played_games = user_recs.sort_values(by='hours',  
ascending=False).head(top_k_precision)  
    user_top_played_game_ids = user_top_played_games['app_id'].tolist()  
  
    recommended_game_ids =  
games_df[games_df['title'].isin(recommended_games)][ 'app_id'].tolist()  
  
    hits = 0  
    for game_id in recommended_game_ids:  
        if game_id in user_top_played_game_ids:  
            hits += 1  
  
    if not recommended_games:  
        return 0.0  
  
    precision = hits / len(recommended_games)  
    return precision
```

Nilai evaluasi akhir ialah mean presisi semua rekomendasi *video game* untuk 100 sampel pengguna.

3.3. Hasil

Berikut adalah hasil pengujian sistem rekomendasi *video game* dengan *collaborative filtering* dengan cara menginput sampel 100 pengguna.

```
[ ] print("Mean precision rekomendasi video game 100 pengguna ialah ", total_precision/100)
Mean precision rekomendasi video game 100 pengguna ialah 0.881166666666667
```

Gambar 3. Hasil Evaluasi Sistem Rekomendasi

Sistem rekomendasi yang diimplementasikan menggunakan algoritma *k-Nearest Neighbors (kNN)* dengan metrik *cosine similarity* berhasil mencapai tingkat presisi yang sangat tinggi, yaitu **88.12%** berdasarkan pengujian terhadap 100 pengguna. Presisi dihitung dengan membandingkan game yang direkomendasikan oleh sistem dengan *top 20* game yang paling banyak dimainkan oleh pengguna target. Hasil ini menunjukkan bahwa sebagian besar rekomendasi yang diberikan benar-benar sesuai dengan preferensi pengguna, karena termasuk dalam daftar game yang sering mereka mainkan. Tingginya nilai presisi mengindikasikan bahwa pendekatan *collaborative filtering* berbasis kemiripan pengguna (*user-based*) efektif dalam memprediksi game yang relevan, terutama ketika data interaksi (dalam hal ini, durasi bermain) mampu mencerminkan pola preferensi pengguna dengan baik. Namun, perlu diperhatikan bahwa evaluasi ini hanya mempertimbangkan *precision* tanpa melihat metrik lain seperti *recall* atau *diversity*, yang juga penting untuk menilai kualitas rekomendasi secara lebih komprehensif. Selain itu, hasil yang baik ini mungkin bergantung pada karakteristik dataset, seperti kepadatan matriks interaksi atau distribusi durasi bermain, sehingga perlu diuji lebih lanjut pada skenario yang lebih beragam.

4. Kesimpulan

Berdasarkan penelitian yang dilakukan, sistem rekomendasi *video game* dengan *Collaborative Filtering* berbasis *playtime* terbukti efektif dalam memberikan rekomendasi yang sesuai dengan preferensi pengguna. Dengan memanfaatkan data *playtime* sebagai indikator interaksi, sistem ini berhasil mencapai presisi sebesar 88,12% pada pengujian terhadap 100 pengguna, menunjukkan bahwa mayoritas rekomendasi yang diberikan termasuk dalam *top 20* game yang paling banyak dimainkan oleh pengguna target. Hasil ini sejalan dengan penelitian sebelumnya yang menggunakan Content-Based Filtering [3], namun pendekatan berbasis kolaborasi ini menawarkan keunggulan dalam memanfaatkan pola perilaku pengguna lain yang memiliki preferensi mirip.

Meskipun demikian, evaluasi ini hanya berfokus pada presisi tanpa mempertimbangkan metrik lain seperti *recall* atau *diversity*, yang dapat memberikan gambaran lebih komprehensif tentang kualitas rekomendasi. Selain itu, performa sistem sangat bergantung pada kualitas dan distribusi data, sehingga pengujian lebih lanjut pada dataset yang lebih beragam diperlukan untuk memastikan konsistensi hasil. Ke depannya, pengembangan sistem dapat ditingkatkan dengan menggabungkan *hybrid recommendation* (menggabungkan *Collaborative dan Content-Based Filtering*) atau memperhitungkan faktor tambahan seperti rating dan ulasan pengguna untuk meningkatkan akurasi dan variasi rekomendasi. Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan *Collaborative Filtering* berbasis *playtime* merupakan solusi yang efektif untuk membantu pengguna menemukan game yang relevan di tengah banyaknya pilihan yang tersedia.

Daftar Pustaka

- [1] Statista.com, 2025, "Video Gaming & eSports," URL: <https://www.statista.com/markets/417/topic/478/video-gaming-esports> [diakses: 30 Juni 2025]
- [2] H. Jaya, H. Winata, I. Zulkarnain, D. Nofriansyah, and P. S. Ramadhan, 2023, "Smart Recommendation System Pemilihan Lokasi Cabang Baru Herman Pangkas Menggunakan

- Metode Vikor,” *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, 22. 1. 10.53513/jis. v22i1.7460.
- [3] D. A. P. D. Pramesti and I. W. Santiyasa, 2022, “Penerapan Metode Content-Based Filtering dalam Sistem Rekomendasi Video Game,” *JNATIA (Jurnal Nasional Teknologi Informasi dan Aplikasinya)*, vol. 1, no. 1, pp. 229-234, Nov. 2022.
- [4] D. V. Gunawan and I. K. A. Mogi, 2024, “Sistem Rekomendasi Produk pada E-Commerce Menggunakan Metode User-Based Collaborative Filtering,” *JNATIA (Jurnal Nasional Teknologi Informasi dan Aplikasinya)*, Vol. 2, No. 4, pp. 753-760, Agu. 2024.
- [5] S. N. Bakri and L. S. Harahap, 2025, “Analisis klasifikasi Algoritma K-Nearest Neighbor (K-NN) pada struktur Daerah di Kota Medan,” *Jurnal Ilmu Komputer dan Sistem Informasi (JIRSI)*, Vol. 4, No. 2, pp. 182-193, Mei 2025