

Sistem Rekomendasi Game dengan Metode K-Nearest Neighbor (KNN)

I Putu Marcel Wiguna^{a1}, Ida Bagus Gede Dwidasmara^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Udayana, Bali
Jln. Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, 08261, Bali, Indonesia
¹marcelwiguna789@gmail.com
²dwidasmara@unud.ac.id

Abstract

The rapid growth of the gaming industry has resulted in an overwhelming number of game titles available to users. However, the abundance of choices makes it challenging for users to find games that match their preferences and interests. To address this issue, this research paper focuses on the development of a game recommendation system. The goal is to create an effective system that assists users in discovering games that align with their tastes and enhances their gaming experience. In this study, the K-Nearest Neighbor (KNN) method is employed as the underlying algorithm for the game recommendation system. The KNN method is a popular machine learning technique known for its ability to classify data based on similarities. This allows the system to recommend games that are likely to be of interest to users based on their preferences and the characteristics of games they have previously enjoyed. This research contributes to the field by showcasing the potential of the K-Nearest Neighbor (KNN) method in developing an efficient game recommendation system. The system's capability to assist users in discovering engaging games tailored to their interests has implications for improving user experience and driving game sales.

Keywords: Game, Recommendation System, KNN

1. Pendahuluan

Perkembangan internet yang sangat pesat telah membawa banyak perubahan dalam perilaku bisnis dan ekonomi masyarakat. Saat ini, banyak situs-situs e-commerce yang sudah dilengkapi kemampuan untuk memberikan rekomendasi sesuai dengan preferensi pelanggan, dengan tujuan supaya pelanggan tertarik untuk membeli produk yang direkomendasikan, contohnya Amazon.com, CDNow, Reel.com, dsb. Salah satu jenis situs yang saat ini semakin marak dan banyak peminatnya adalah situs-situs penyedia game. Situs-situs ini biasanya menyediakan berbagai jenis game untuk komputer maupun video game [1].

Sistem rekomendasi merupakan sistem yang dapat membantu pengguna dalam mengidentifikasi hal-hal yang mungkin diminati oleh pengguna tersebut. Sistem ini dibuat untuk memprediksikan produk yang diminati kemudian memberikan rekomendasi kepada pengguna berdasarkan banyak faktor. Sistem rekomendasi memproses sejumlah besar informasi pengguna, preferensi, dan minat pengguna. Data tersebut akan digunakan untuk menemukan kecocokan dan menghitung kesamaan antara pengguna dan item. Sistem rekomendasi dapat digunakan dalam berbagai hal, misalnya menemukan item yang diminati pengguna, membantu penyedia barang dalam mengirimkan barang kepada pengguna yang tepat, dan lain-lain [2].

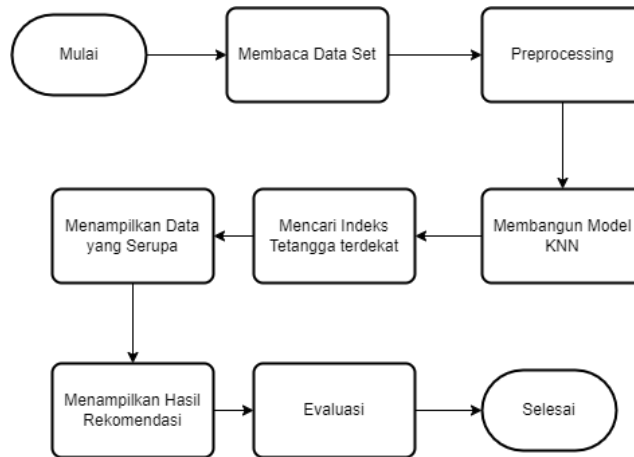
Dalam sistem rekomendasi terdapat beberapa metode yang sering digunakan yaitu Collaborative Filtering, Content Based, dan Hybrid. Collaborative Filtering menggunakan riwayat pemilihan atau riwayat penilaian sebagai dasar untuk menentukan rekomendasi. Content Based menggunakan kesamaan produk untuk ditawarkan kepada pembeli. Sedangkan metode Hybrid menggabungkan metode dua atau lebih metode untuk menghasilkan rekomendasi yang lebih baik [3]. Dalam penelitian ini peneliti akan menggunakan K-Nearest Neighbor karena KNN dapat

digunakan untuk menangani variasi data yang kompleks atau tidak terstruktur. Dalam sistem rekomendasi game, ini berarti KNN dapat memberikan rekomendasi yang lebih relevan dan beragam, berdasarkan kesamaan fitur-fitur numerik antara game-game.

2. Metode Penelitian

a. Konsep Dari Sistem Yang Dibangun

Sebelum dilakukannya tahap penelitian, dibuat terlebih dahulu perancangan dari sistem yang akan dibangun yang dimana akan memudahkan untuk memahami alur dari sistem itu sendiri.



Gambar 1. Flowchart Sistem

b. Dataset

Dataset yang akan digunakan dalam penelitian ini didapatkan dari situs Kaggle, yang dimana dataset ini berformat CSV. Dataset yang digunakan berisi sejumlah data game, platform, tahun rilis dsb.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Name	Platform	Year_of_R	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sale	Global_Sal	Critic_Scor	Critic_Cou	User_Scor	User_Cou	Developer	Rating
2	Wii Sports	Wii	2006	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76	51	8	322	Nintendo	E
3	Super Mar	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24						
4	Mario Kart	Wii	2008	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52	82	73	8.3	709	Nintendo	E
5	Wii Sports	Wii	2009	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77	80	73	8	192	Nintendo	E
6	Pokemon	GB	1996	Role-Playii	Nintendo	11.27	8.89	10.22	1	31.37						
7	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58	30.26						
8	New Super	DS	2006	Platform	Nintendo	11.28	9.14	6.5	2.88	29.8	89	65	8.5	431	Nintendo	E
9	Wii Play	Wii	2006	Misc	Nintendo	13.96	9.18	2.93	2.84	28.92	58	41	6.6	129	Nintendo	E
10	New Super	Wii	2009	Platform	Nintendo	14.44	6.94	4.7	2.24	28.32	87	80	8.4	594	Nintendo	E
11	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31						
12	Nintendog	DS	2005	Simulation	Nintendo	9.05	10.95	1.93	2.74	24.67						
13	Mario Kart	DS	2005	Racing	Nintendo	9.71	7.47	4.13	1.9	23.21	91	64	8.6	464	Nintendo	E
14	Pokemon	GB	1999	Role-Playii	Nintendo	9	6.18	7.2	0.71	23.1						
15	Wii Fit	Wii	2007	Sports	Nintendo	8.92	8.03	3.6	2.15	22.7	80	63	7.7	146	Nintendo	E
16	Kinect Adv	X360	2010	Misc	Microsoft	15	4.89	0.24	1.69	21.81	61	45	6.3	106	Good Scie	E
17	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.01	8.49	2.53	1.77	21.79	80	33	7.4	52	Nintendo	E
18	Grand The	PS3	2013	Action	Take-Two	7.02	9.09	0.98	3.96	21.04	97	50	8.2	3994	Rockstar	N M
19	Grand The	PS2	2004	Action	Take-Two	9.43	0.4	0.41	10.57	20.81	95	80	9	1588	Rockstar	N M
20	Super Mar	SNES	1990	Platform	Nintendo	12.78	3.75	3.54	0.55	20.61						
21	Brain Age:	DS	2005	Misc	Nintendo	4.74	9.2	4.16	2.04	20.15	77	58	7.9	50	Nintendo	E
22	Pokemon	DS	2006	Role-Playii	Nintendo	6.38	4.46	6.04	1.36	18.25						
23	Super Mar	GB	1989	Platform	Nintendo	10.83	2.71	4.18	0.42	18.14						
24	Super Mar	NES	1988	Platform	Nintendo	9.54	3.44	3.84	0.46	17.28						
25	Grand The	X360	2013	Action	Take-Two	9.66	5.14	0.06	1.41	16.27	97	58	8.1	3711	Rockstar	N M
26	Grand The	PS2	2002	Action	Take-Two	8.41	5.49	0.47	1.78	16.15	95	62	8.7	730	Rockstar	N M
27	Pokemon	GBA	2002	Role-Playii	Nintendo	6.06	3.9	5.38	0.5	15.85						

c. K-Nearest Neighbor

K-Nearest Neighbour (K-NN) merupakan metode untuk melakukan klasifikasi terhadap objek baru berdasarkan (k) tetangga terdekatnya. K-NN termasuk algoritma supervised learning, dimana hasil dari query instance yang baru, diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN, kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi [4]. Metode KNN digunakan karena memiliki beberapa kelebihan, diantaranya yaitu dapat menghasilkan data yang lebih akurat dan efektif apabila memiliki training data yang cukup besar. Namun, metode ini juga memiliki beberapa kekurangan, seperti biaya komputasi yang cukup tinggi karena diperlukan perhitungan jarak query instance pada keseluruhan training sample [5]. Persamaan berikut ini adalah persamaan :

Distance functions

Euclidean	$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$
Manhattan	$\sum_{i=1}^k x_i - y_i $
Minkowski	$\left(\sum_{i=1}^k (x_i - y_i ^q) \right)^{1/q}$

Gambar 2. Persamaan Pada Metode KNN

3. Hasil dan Pembahasan

Adapun tahapan-tahapan yang dilakukan sampai mendapatkan hasil dari penelitian adalah sebagai berikut.

3.1. Preprocessing

Pada tahap preprocessing ini, langkah-langkah seperti mengambil fitur-fitur yang relevan, mengubah nilai string menjadi NaN, dan menghapus baris dengan nilai NaN adalah langkah-langkah yang umum dilakukan untuk membersihkan dan mempersiapkan data sebelum digunakan dalam analisis lebih lanjut. Selain itu, normalisasi pada juga termasuk dalam tahap ini. Normalisasi dilakukan untuk mengubah rentang nilai pada fitur-fitur numerik agar seragam atau memiliki skala yang sama. Dalam kasus ini, normalisasi dilakukan dengan mengubah nilai-nilai fitur numerik ke dalam rentang antara 0 dan 1.

```
# Normalisasi fitur-fitur numerik
dataset[features] = (dataset[features] - dataset[features].min()) / (dataset[features].max() - dataset[features].min())

# Menampilkan data setelah pengolahan
print("Data setelah pengolahan:")
print(dataset.head())
```

Data setelah pengolahan:

	Name	Platform	Year_of_Release	Genre	Publisher	\
0	Wii Sports	Wii	0.677419	Sports	Nintendo	
2	Mario Kart Wii	Wii	0.741935	Racing	Nintendo	
3	Wii Sports Resort	Wii	0.774194	Sports	Nintendo	
6	New Super Mario Bros.	DS	0.677419	Platform	Nintendo	
7	Wii Play	Wii	0.677419	Misc	Nintendo	

	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	\
0	41.36	28.96	3.77	8.45	82.53	0.741176	
2	15.68	12.76	3.79	3.29	35.52	0.811765	
3	15.61	10.93	3.28	2.95	32.77	0.788235	
6	11.28	9.14	6.50	2.88	29.80	0.894118	
7	13.96	9.18	2.93	2.84	28.92	0.529412	

	Critic_Count	User_Score	User_Count	Developer	Rating	\
0	0.436364	0.824176	0.029828	Nintendo	E	
2	0.636364	0.857143	0.066129	Nintendo	E	
3	0.636364	0.824176	0.017634	Nintendo	E	
6	0.563636	0.879121	0.040053	Nintendo	E	
7	0.345455	0.670330	0.011725	Nintendo	E	

Gambar 3. Hasil Preprocessing

3.2. Tahap Pembangunan Model

Pada tahap ini, kita menggunakan dataset yang telah diolah dan dinormalisasi fitur-fiturnya. Setelah itu, kita membuat objek KNN dengan menginisialisasi jumlah tetangga ($n_neighbors$) yang ingin kita pertimbangkan saat mencari tetangga terdekat, algoritma yang digunakan, dan metrik jarak yang digunakan. Setelah objek KNN dibuat, kita melakukan proses pembelajaran dengan memasukkan dataset yang telah dipreproses ke dalam algoritma KNN. Proses pembelajaran ini melibatkan menemukan tetangga terdekat untuk setiap titik data dalam dataset berdasarkan jarak yang ditentukan. Algoritma KNN akan mencatat tetangga-tetangga ini sebagai referensi untuk memberikan rekomendasi di kemudian hari. Selama tahap pembangunan model, KNN akan menghitung jarak antara setiap titik data dengan tetangga terdekatnya, yang akan digunakan sebagai dasar perhitungan kesamaan antara data. Semakin dekat jarak antara dua data, semakin mirip fitur-fitur numeriknya. Setelah proses pembelajaran selesai, model KNN akan siap digunakan untuk memberikan rekomendasi game. Model ini memiliki pengetahuan tentang tetangga terdekat setiap titik data dalam dataset, dan berdasarkan kesamaan fitur-fiturnya, dapat merekomendasikan game-game serupa untuk game yang diberikan sebagai input.

3.3. Tahap Evaluasi

Pada tahap ini akan dilakukan kinerja sistem yang telah dibangun yang dimana sangat penting bagi peneliti jika ingin meningkatkan kinerja model yang telah dibangun. Dalam tahap ini matrik pengujian yang digunakan Root Mean Squared Error (ERMS) yang dimana dipergunakan dalam mengukur tingkat akurasi dari hasil rekomendasi oleh model dari sistem yang telah dibangun. Berikut adalah hasil dari rekomendasi sistem sekaligus pengukuran ERMS terhadap model yang telah dibangun oleh peneliti.

```
▶ game_name = input("Masukkan nama game: ")  
  
# Mencari rekomendasi game  
recommended_games = recommend_games(game_name, dataset, knn)  
  
# Menampilkan rekomendasi game  
print("Game yang direkomendasikan:")  
print(recommended_games)  
  
# Menghitung RMSE untuk game input  
rmse = calculate_rmse(game_name, dataset, knn)  
print("RMSE:", rmse)  
  
↳ Masukkan nama game: Wii Play  
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: User  
warnings.warn(  
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: User  
warnings.warn(  
Game yang direkomendasikan:  
4020 Tony Hawk's Project 8  
4170 Viva Pinata: Trouble in Paradise  
1810 Mirror's Edge  
644 Tom Clancy's Rainbow Six: Vegas 2  
4450 The Darkness  
Name: Name, dtype: object  
RMSE: 0.14326188024767245
```

Gambar 4. Hasil Evaluasi

Hasil RMSE yang diperoleh adalah 0.14326188024767245, maka itu menunjukkan bahwa kesalahan rata-rata dari model rekomendasi game menggunakan metode K-Nearest Neighbors (KNN) sebesar 0.14326188024767245. Yang dimana semakin rendah nilai RMSE, semakin baik performa model. Dalam konteks ini, nilai RMSE yang relatif rendah menunjukkan bahwa model KNN mampu memberikan rekomendasi game yang cukup akurat.

4. Kesimpulan

Berdasarkan dari hasil penelitian yang telah dilakukan dapat disimpulkan model yang telah dibangun telah dapat memberikan rekomendasi yang cukup memuaskan berdasarkan hasil evaluasi yang dilakukan, namun peneliti juga merasa untuk teknik evaluasi yang digunakan ada yang dapat lebih bisa dipercaya disbanding RMSE dan juga untuk penelitian kedepannya peneliti menyarankan untuk penggunaan dataset yang lebih banyak lagi dan tentunya beragam agar dapat menghasilkan model yang lebih optimal.

Daftar Pustaka

- [1] H. Maharani, "Rancangan Sistem Rekomendasi Game Dengan Model-Based Collaboration Filtering."
- [2] H. A. Sheeren, L. W. Santoso, and J. Andjarwirawan, "Sistem Rekomendasi Pembelian Laptop dengan K-Nearest Neighbor (KNN)."
- [3] C. Saha and D. Prasetya, "Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor," vol. 4, no. 3, pp. 2355–7699, 2017.
- [4] C. A. Rahardja and H. Agung, "Rahardja, Implementasi Algoritma K-Nearest Neighbor Pada Website Rekomendasi Laptop 75 Implementasi Algoritma K-Nearest Neighbor Pada Website Rekomendasi Laptop."
- [5] P. Pangestu and R. Setyadi, "Penerapan Metode K-Nearest Neighbor Untuk Pemilihan Rekomendasi Game FPS Pada Aplikasi Google Play Store," *Journal of Information*

System Research (JOSH), vol. 4, no. 2, pp. 742–747, Jan. 2023, doi: 10.47065/josh.v4i2.3006.