

Jurnal.docx

by Check Turnitin

Submission date: 16-Dec-2024 10:42PM (UTC-0800)

Submission ID: 2467077658

File name: Jurnal.docx (621.99K)

Word count: 2783

Character count: 18029

Klasterisasi Data Top Rated Games On Steam Dengan Algoritma K-Means dan Metode Elbow Di Rapidminer

Penulis I Ikhlashul Abi Monhestisaura⁽¹⁾

⁽¹⁾ Jurusan Informatika, Universitas Pancasakti Tegal, Indonesia
Email Penulis: Ikhlashulabi13@gmail.com

Abstrak

Penelitian ini membahas penerapan algoritma K-Means untuk klasterisasi data game dengan rating tertinggi di Steam menggunakan metode elbow. Proses klasterisasi dilakukan di RapidMiner dengan memanfaatkan berbagai operator, seperti Select Attributes, Multiply, K-Means (dengan variasi cluster 2, 3, 5, dan 7), serta Cluster Distance Performance untuk mengevaluasi hasil. Berdasarkan analisis, jumlah cluster terbaik yang dihasilkan adalah 3, sesuai dengan titik siku grafik elbow. Penelitian ini bertujuan membantu pengguna dalam mengelompokkan game untuk mengetahui berbagai jenis review setiap game, seperti Overwhelmingly Positive, Most Positive atau Very Positive.

21

Kata Kunci: Klasterisasi, K-Means, Metode Elbow, Rapidminer, Game Steam

Pendahuluan

Dalam era digital yang semakin berkembang, Kemampuan untuk mengumpulkan, menganalisis, dan menggunakan data dengan bijak akan menjadi kunci kesuksesan di berbagai bidang, data menjadi salah satu aset terpenting dalam berbagai bidang, termasuk industri game [1]. Platform distribusi game digital seperti Steam menghadirkan ribuan permainan yang mendapat penilaian dari pengguna di seluruh dunia. Dengan bertambahnya jumlah data, analisis yang tepat diperlukan untuk memahami pola perilaku pengguna dan mengevaluasi kualitas game. Salah satu metode yang efektif untuk memahami struktur data besar adalah klasterisasi, yang bertujuan mengelompokkan data berdasarkan karakteristik tertentu. Clustering merupakan teknik untuk mengelompokkan data dan membentuk kelas objek-objek yang memiliki kemiripan. Dalam penelitian ini peneliti juga menggunakan algoritma *K-means*. Kelebihan dari algoritma *K-means* clustering adalah metode partisi yang paling efisien dan telah digunakan dalam berbagai aplikasi karena kesederhanaan dan dapat mencapai konvergensi yang cepat [2].

Algoritma *K-Means* menjadi pilihan utama karena efisiensinya dalam menangani data besar dan kemampuannya menghasilkan klaster yang bermakna. Dan *K-means* juga algoritma paling mudah digunakan untuk klasterisasi pada pemula. Namun, dalam implementasi algoritma K-Means ada satu tantangan utama yang harus diperhatikan yaitu menentukan jumlah cluster yang optimal. Jika jumlah cluster tidak tepat, maka hasil klasterisasi dapat menjadi tidak representatif atau sulit diinterpretasikan. Untuk mengatasi masalah ini, metode elbow digunakan sebagai alat evaluasi yang membantu menentukan jumlah cluster berdasarkan rata-rata jarak titik centroid pada cluster tersebut. Grafik elbow memberikan visualisasi yang jelas untuk menemukan jumlah cluster terbaik yang sesuai dengan data yang dianalisis. Dimana titik paling siku yang optimal dari jarak titik centroid.

Penelitian ini difokuskan pada dataset "Top Rated Games on Steam," yang peneliti kumpulkan dari *database steam* game yang mencakup atribut-atribut seperti id, nama game, jumlah ulasan positif, jumlah ulasan negatif, total ulasan dan rating game. Dataset ini dipilih karena dapat memberikan wawasan berharga tentang segmentasi game berdasarkan popularitas dan penilaian pengguna. Menggunakan Rapidminer, penelitian ini memanfaatkan operator seperti Select Attributes untuk pemilihan data penting, Multiply untuk memperbanyak dataset, dan K-Means dengan variasi

jumlah cluster. Evaluasi dilakukan dengan menggunakan operator *Cluster Distance Performance* untuk menganalisis kualitas klusterisasi. Setelah dianalisis, ambil kesimpulan mana jumlah cluster yang optimal menggunakan metode elbow.

Tujuan dari penelitian ini adalah untuk menerapkan algoritma *K-means* dalam mengelompokkan data game pada platform Steam, menentukan jumlah cluster terbaik atau optimal menggunakan metode elbow, serta mengevaluasi hasil klusterisasi berdasarkan interpretasi data. Dengan menemukan jumlah cluster optimal, penelitian ini diharapkan dapat membantu para pengembang game, peneliti, atau pemasar untuk memahami segmentasi pasar game secara lebih mendalam. Serta para *gamers* bisa mengetahui game pada platform steam yang memiliki penilaian baik berdasarkan ulasan positif dan negatif, yang mana penting dalam pembelian game.

Manfaat penelitian ini bersifat teoritis dan praktis. Secara teoritis, penelitian ini memberikan kontribusi dalam pengembangan metodologi analisis data menggunakan algoritma *K-means* dan metode elbow yang di implementasikan pada software rapidminer, khususnya dalam aplikasi pada dataset industri kreatif seperti game. Secara praktis, hasil klusterisasi dapat dimanfaatkan untuk mengembangkan strategi pemasaran, rekomendasi game kepada pengguna, atau bahkan merancang game baru atau *sequel* dari game sebelumnya berdasarkan segmentasi yang telah teridentifikasi dalam pengelompokan ulasan rating tertinggi.

Dengan demikian, penelitian ini tidak hanya relevan dalam konteks akademik tetapi juga memberikan dampak signifikan pada industri game yang menjadi salah satu sektor yang berkembang pesat dan berpotensi besar untuk berkontribusi terhadap perekonomian nasional, terutama dalam memahami kebutuhan dan preferensi pengguna. Penelitian lebih lanjut dapat memperluas cakupan analisis dengan memasukkan atribut tambahan seperti harga, genre, atau waktu bermain untuk menghasilkan klusterisasi yang lebih komprehensif.

Lansasan Teori

A. Data Mining

Data mining adalah proses mengekstraksi data dan informasi penting dari kumpulan data yang lebih besar dengan menggunakan matematika, berbagai teknik statistik, dan AI. Clustering, analisis regresi, dan text mining adalah beberapa teknik yang digunakan dalam data mining. Klusterisasi, yang memungkinkan pengelompokan data berdasarkan fitur yang sebanding, adalah salah satu teknik yang umum digunakan dalam pengolahan data. Data mining sangat berguna dalam banyak bidang, seperti pemasaran, kesehatan, dan hiburan, termasuk analisis data game [3, p. 2].

B. Algoritma K-Means

K-Means merupakan metode Clustering yang memiliki kinerja dalam mengelompokkan dasar data item dalam bentuk fitur yang baru pada grup C. Nilai C pada metode ini merupakan nilai ketentuan kelas yang akan terbentuk berdasarkan pada nilai yang di masukan pada nilai C itu sendiri. Penginputan nilai selalu dalam bentuk Integer yang positif. Pengelompokan pada metode k-means akan dilakukan dengan cara mengurangi jumlah total kuadrat jarak antara Cluster centroid dan data. Ada beberapa tahapan yang di terapkan pada algoritma k-means yaitu sebagai berikut:

- Menentukan nilai C yang akan dibuat
- Menentukan nilai Centroid secara acak
- Menghitung jarak pada setiap data ke centroid terdekat
- Implementasikan dengan menggunakan rumus jarak euclidean

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan :

- $\sum_{i=1}^n$, $i=1$ bermakna bahwa perhitungan penjumlahan dimulai dari atribut pertama hingga n , n merupakan jumlah total dari atribut data yang akan dihitung.
- $(xi - yi)^2$, xi merupakan data awal dari atribut pertama dan nilai dari yi merupakan hasil rata rata dari semua data cluster pada iterasi ke-2 dan seterusnya.
- e. Perbarui centroid cluster

$$C_j = \frac{1}{n_j} \sum_{i=1}^{n_j} X_i \quad (2)$$

Keterangan :

- n_j , merepresentasikan jumlah data dari j (cluster) yang sudah di alokasikan sebelumnya.
- $\sum_{i=1}^{n_j} x_i$, $i=1$ bermakna bahwa perhitungan penjumlahan dimulai dari atribut pertama hingga n_j , n_j merupakan jumlah total dari atribut data cluster yang akan dihitung sesuai dengan posisi atribut ke i .
- f. Ulangi iterasi pada b dan c hingga tidak ada perubahan yang signifikan dalam klasterisasi data [3, p. 3].

C. Metode Elbow

Metode Elbow merupakan metode pendukung yang dapat menentukan nilai C yang optimal pada penerapan metode K-Means. Penerapan metode ini memiliki fokus pada persentase varian sebagai fungsi dari jumlah Cluster. Untuk mencari nilai C yang optimal maka nilai C akan dicek satu persatu dan akan dicatat nilai SSE (*Sum Square Error*). di mana jumlah cluster optimal ditandai oleh titik siku pada grafik. Titik ini menunjukkan ketika penurunan SSE mulai melambat, yang berarti tambahan cluster tidak memberikan peningkatan signifikan dalam kualitas klasterisasi. Dengan metode elbow, proses klasterisasi menjadi lebih terarah dan menghasilkan segmentasi data yang lebih representatif.

$$SSE = \sum_{k=1}^k \sum_{xi \in Sk} \|Xi - Ck\|^2 \quad (3)$$

Nilai SSE adalah jumlah rata-rata dari Jarak Euclidean pada setiap titik terhadap Centroid. Pada diagram bila nilai turun secara signifikan dan membuat lekukan dari garis pada titik awal dan setelahnya maka nilai dari C telah di temukan [4].

D. Rapidminer

RapidMiner adalah perangkat lunak berbasis Java yang dirancang untuk mendukung proses data mining dengan antarmuka pengguna yang ramah. Platform ini menyediakan alat untuk klasterisasi, klasifikasi, dan analisis regresi, serta mendukung berbagai format file seperti .xls dan .csv. RapidMiner dirancang untuk mendukung semua langkah dalam ilmu data, mulai dari persiapan data hingga implementasi pembelajaran mesin dan distribusi model prediktif. Dengan lebih dari 625.000 pengguna profesional, RapidMiner digunakan secara luas dalam bisnis, pendidikan, dan penelitian untuk meningkatkan efisiensi analisis data. Platform ini menawarkan

tiga perspektif utama—Welcome, Design, dan Result—untuk memudahkan pengguna mendesain, menjalankan, dan mengevaluasi proses analisis [5].

E. Video Game

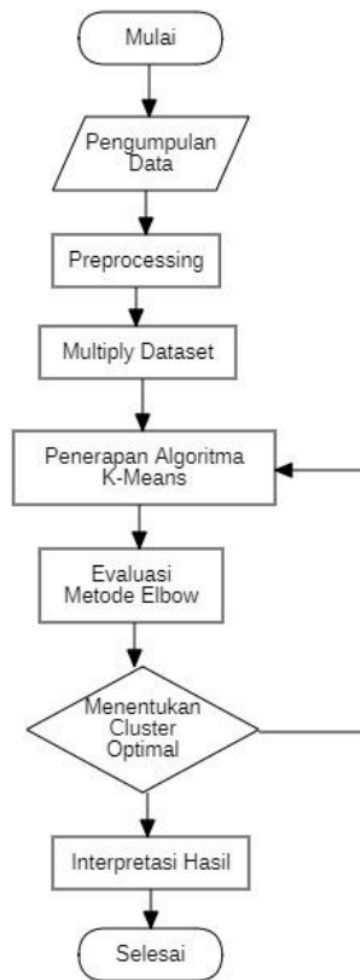
Video Game adalah sebuah bentuk permainan elektronik berupa teks maupun gambar, yang melibatkan interaksi antara perangkat lunak permainan, orang yang memainkannya, dan dijabatani oleh perangkat keras pengolah permainan tersebut. Perangkat lunak game akan memberikan output berupa gambar atau teks yang ditampilkan melalui media (televisi, komputer, telepon seluler, dll.), kemudian pemain memberikan input berupa perintah yang disalurkan melalui perangkat keras permainan tersebut untuk ditampilkan kembali pada media. Game mencakup berbagai genre seperti aksi, strategi, petualangan, dan simulasi, yang dirancang untuk memenuhi kebutuhan pengguna yang beragam [6].

F. Steam

Steam merupakan salah satu platform e-commerce produk digital yang diciptakan oleh Valve Corporation. Valve pertama kali meluncurkan platform Steam pada tahun 2003 yang akhirnya berkembang pesat hingga kini. Platform ini menawarkan layanan Digital Right Management (DRM), permainan video multi-pemain, jejaring sosial, dan streaming video. Steam menjual game berbasis komputer. Selain menyediakan toko untuk pembelian game, Steam juga memberikan layanan komunitas, ulasan pengguna, dan fitur statistik seperti jumlah pemain aktif. Data dari Steam sering kali digunakan untuk menganalisis popularitas game, perilaku pemain, dan tingkat kepuasan pengguna. Dengan data yang kaya dan terstruktur, Steam menjadi sumber yang ideal untuk penelitian di bidang analisis data game [7].

Metode Penelitian

Penelitian ini dilakukan menggunakan dataset Top Rated Games on Steam untuk menentukan cluster terbaik dengan algoritma K-Means dan metode elbow. Proses penelitian dimulai dari pengumpulan dataset, preprocessing data, *multiply* data, penerapan algoritma K-Means, evaluasi metode *elbow*, hingga hasil klusterisasi. Seluruh langkah dilakukan menggunakan perangkat lunak RapidMiner yang mendukung metode drag-and-drop untuk memudahkan implementasi. Berikut adalah gambar flowchart dari tahapan proses tersebut



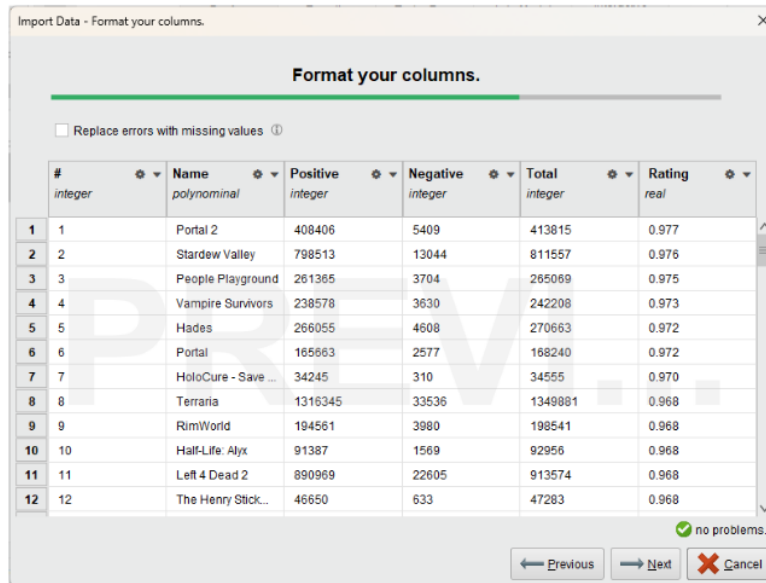
Gambar 1. Diagram alur penelitian

Hasil Penelitian Dan Pembahasan

Penelitian bertujuan untuk menentukan jumlah cluster optimal pada dataset *Top Rated Games on Steam* menggunakan algoritma K-Means dan metode elbow di RapidMiner. Hasil penelitian diperoleh dari serangkaian pengumpulan dataset hingga analisis klasterisasi dengan variasi jumlah cluster (2, 3, 5, dan 7). Hasil setiap analisis kemudian dievaluasi menggunakan metode elbow.

A. Pengumpulan dataset

Dataset diambil dari sumber terpercaya yang mencakup informasi seperti nama game, positive, negative, total, dan rating. Data tersebut akan di proses klasterisasi setelah melewati Langkah selanjutnya



Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

#	Name	Positive	Negative	Total	Rating
1	Portal 2	408406	5409	413815	0.977
2	Stardew Valley	798513	13044	811557	0.976
3	People Playground	261365	3704	265069	0.975
4	Vampire Survivors	238578	3630	242208	0.973
5	Hades	266055	4608	270663	0.972
6	Portal	165663	2577	168240	0.972
7	HaloCure - Save ...	34245	310	34555	0.970
8	Terraria	1316345	33536	1349881	0.968
9	RimWorld	194561	3980	198541	0.968
10	Half-Life: Alyx	91387	1569	92956	0.968
11	Left 4 Dead 2	890969	22605	913574	0.968
12	The Henry Stick...	46650	633	47283	0.968

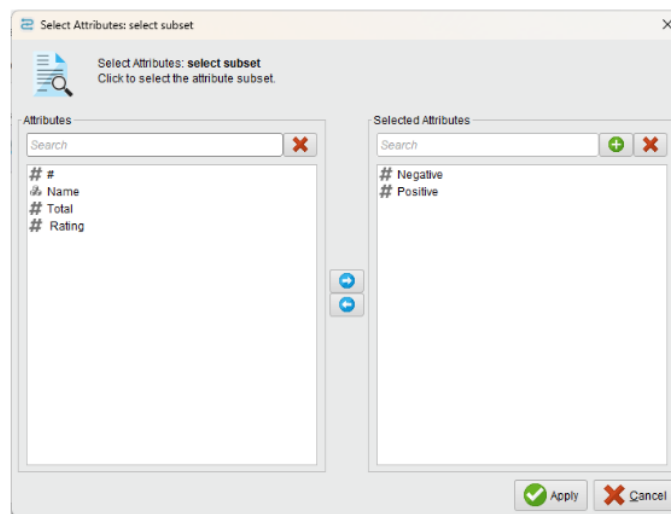
no problems.

Previous Next Cancel

Gambar 2. Dataset *Top Rated Games on Steam*

B. Preprocessing

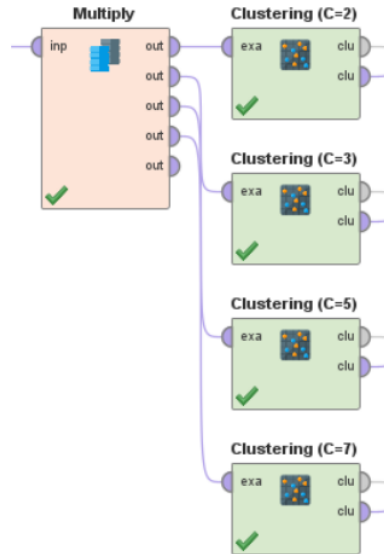
Proses preprocessing dilakukan untuk memastikan kualitas data yang optimal. Tahap ini mencakup seleksi data, dan memilih dua atribut utama, yaitu *positive* dan *negative review*. Tahapan ini dilakukan menggunakan operator Select Attributes di RapidMiner.



Gambar 3. Proses *Select Attribut* data game

C. Multiply data

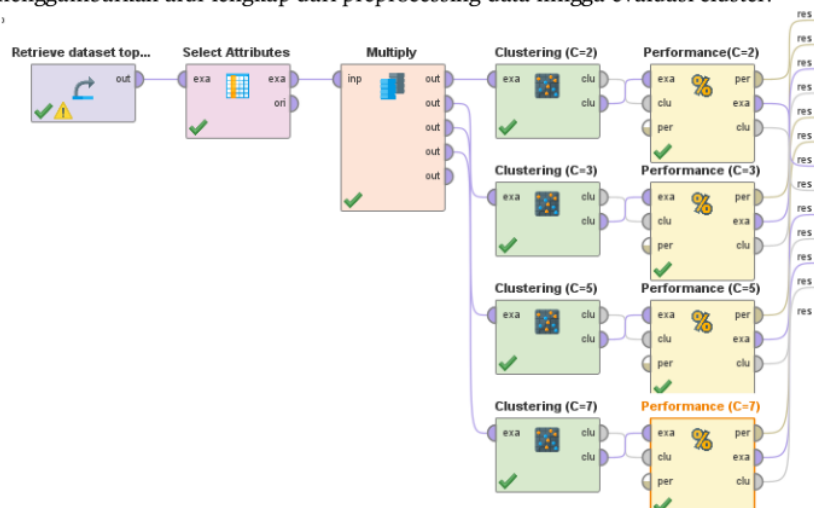
Untuk menguji berbagai jumlah cluster, dataset digandakan menggunakan operator Multiply. Empat salinan dataset dibuat untuk diuji dengan nilai cluster 2, 3, 5, dan 7.



Gambar 4. Multiply data dengan cluster berbeda

12
D. Hasil Penerapan Algoritma K-means

Algoritma K-Means diterapkan pada masing-masing dataset yang digandakan. Setelah itu, Desain proses klusterisasi dibuat dalam bentuk workflow di RapidMiner untuk memastikan langkah-langkah berjalan sistematis dan terstruktur. Proses ini ditampilkan pada Gambar 5, yang menggambarkan alur lengkap dari preprocessing data hingga evaluasi cluster.



Gambar 5. Desain proses klusterisasi

Kemudian proses klusterisasi pada *dataset Top Rated Games on Steam* menghasilkan performa untuk masing-masing cluster berdasarkan analisis rata-rata jarak tiap cluster. Evaluasi ini dilakukan menggunakan operator Cluster Distance Performance di RapidMiner. Nilai *Sum of Squared Errors* (SSE) dihitung untuk setiap jumlah cluster (c) guna menentukan kualitas klusterisasi.



20
Gambar 6. Hasil *performance* cluster 2



Gambar 7. Hasil *performance* cluster 3



Gambar 8. Hasil *performance* cluster 5



Gambar 9. Hasil *performance* cluster 7

Hasil performa untuk masing-masing jumlah cluster yang berdasarkan *avg. within centroid distance* ditunjukkan pada Tabel 1 berikut:

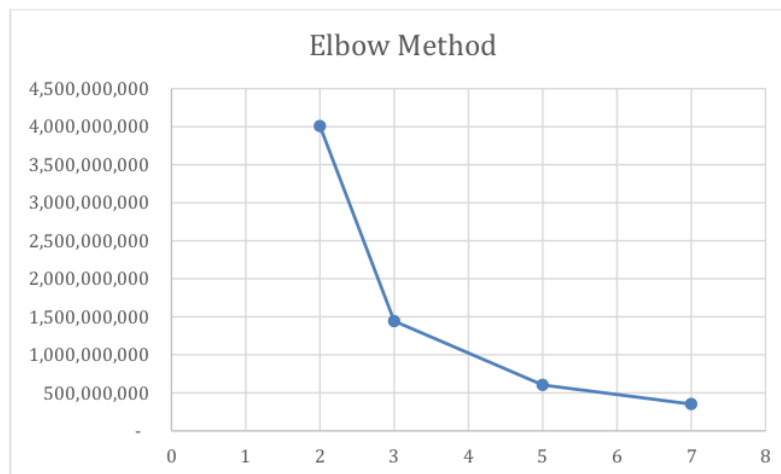
Tabel 1. Hasil pengujian *performance* berdasarkan *avg. within centroid distance*

#	cluster	avg. centroid
1	2	4,007,832,704
2	3	1,443,010,669
3	5	603,224,869
4	7	351,967,320

Dari tabel tersebut, terlihat bahwa nilai SSE cenderung menurun seiring bertambahnya jumlah cluster. Penurunan signifikan terjadi dari cluster 2 ke cluster 3, sedangkan setelah cluster 3, laju penurunan mulai melambat. Dengan itu, penelitian akan menggunakan metode elbow pada Langkah selanjutnya

E. Hasil Evaluasi dengan metode *elbow*

Evaluasi klusterisasi dilakukan menggunakan metode elbow untuk menentukan jumlah cluster optimal pada data Top Rated Games on Steam. Metode ini mengevaluasi jumlah cluster (c) berdasarkan nilai Sum of Squared Errors (SSE). Nilai SSE dihitung untuk setiap cluster yang diuji, yaitu 2, 3, 5, dan 7. Grafik hubungan antara jumlah cluster dan nilai SSE ditampilkan pada Gambar 10.



Gambar 10. Grafik metode elbow berdasarkan tabel 1

Dari grafik yang ditampilkan, nilai SSE terlihat menurun seiring bertambahnya jumlah cluster. Pada cluster 2, nilai SSE masih tinggi, menunjukkan bahwa variasi dalam cluster masih cukup besar. Penurunan signifikan terjadi pada cluster 3, sedangkan setelah itu, laju penurunan SSE mulai melambat. Penurunan yang lambat ini mengindikasikan bahwa jumlah cluster tambahan tidak lagi memberikan peningkatan signifikan terhadap kualitas klusterisasi. Titik siku atau elbow point terlihat jelas pada cluster 3. Pada titik ini, kurva perubahan SSE mulai mendatar, menunjukkan bahwa tiga cluster merupakan jumlah optimal. Hal ini menandakan bahwa jumlah cluster tersebut mampu memberikan keseimbangan antara kehomogenan data dalam cluster dan heterogenan antarcluster.

F. Hasil Cluster optimal

Setelah melalui proses evaluasi menggunakan metode elbow, jumlah cluster optimal yang dipilih adalah 3. Setiap cluster yang terbentuk memiliki karakteristik unik berdasarkan hasil klasterisasi terhadap dataset *Top Rated Games on Steam*. Karakteristik ini menggambarkan distribusi data dan memberikan wawasan tentang persepsi pengguna terhadap game berdasarkan atribut seperti positif dan negative review. Dari total item dibagi beberapa karakteristik berdasarkan cluster nya. Berikut karakteristik dari cluster tersebut

- a) Cluster 0 (Overwhelmingly Positive): Cluster ini mencakup game dengan skor ulasan tertinggi, yang sering kali dinilai sangat positif oleh mayoritas pemain. Game dalam cluster ini biasanya memiliki jumlah pemain yang tinggi serta reputasi yang luar biasa di komunitas pemain Steam.
- b) Cluster 1 (Mostly Positive): Cluster ini berisi game dengan ulasan yang positif secara keseluruhan, namun dengan nilai yang sedikit lebih rendah dibandingkan Cluster 0. Game dalam kelompok ini cenderung memiliki komunitas pemain yang loyal tetapi tidak sepopuler game dalam Cluster 0.
- c) Cluster 2 (Very Positive): Cluster ini terdiri dari game dengan ulasan yang baik namun lebih bervariasi dibandingkan dua cluster lainnya. Game di cluster ini sering kali memiliki basis pemain yang lebih kecil tetapi tetap memiliki daya tarik tertentu bagi segmen pasar tertentu.



Gambar 11. Total data *games on steam*

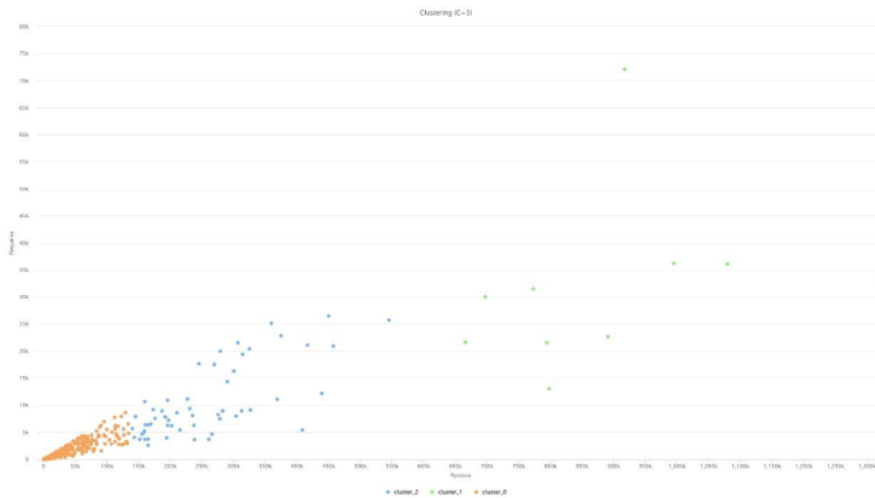
Berikut Data hasil klasterisasi untuk cluster 3 ditampilkan dalam Gambar 10. menampilkan hasil akhir proses klasterisasi dataset *Top Rated Games on Steam* menggunakan algoritma K-Means dengan jumlah cluster (*c*) optimal sebanyak 3. Setiap baris pada tabel merepresentasikan satu instance (game) dalam dataset, yang telah dikelompokkan ke dalam salah satu dari tiga cluster. ini memberikan gambaran yang jelas mengenai bagaimana data terbagi ke dalam tiga cluster berdasarkan ulasan positif dan negatif. Distribusi cluster memperkuat evaluasi bahwa jumlah cluster optimal adalah 3 karena adanya perbedaan karakteristik yang signifikan antarcluster

Row No.	id	cluster	Positive	Negative
1	1	cluster_2	408406	5409
2	2	cluster_1	798513	13044
3	3	cluster_2	261365	3704
4	4	cluster_2	238578	3630
5	5	cluster_2	266055	4608
6	6	cluster_2	165663	2577
7	7	cluster_0	34245	310
8	8	cluster_1	1316345	33536
9	9	cluster_2	194561	3980
10	10	cluster_0	91387	1569
11	11	cluster_1	890969	22605
12	12	cluster_0	46650	633
13	13	cluster_0	27221	282

ExampleSet (1,000 examples,2 special attributes,2 regular attributes)

Gambar 12. Hasil Data klasterisasi dengan cluster 3

Gambar 13. menunjukkan visualisasi scatter/bubble plot dari hasil klasterisasi. Pada visualisasi ini, setiap cluster direpresentasikan dengan warna yang berbeda berdasarkan atribut *Positive* (sumbu X) dan *Negative* (sumbu Y) yang menggambarkan jumlah ulasan positif dan negatif pada game di *Steam* untuk mempermudah interpretasi. Cluster 0, Cluster 1, dan Cluster 2 dapat dengan jelas dilihat dari penyebaran data, dengan jarak antarcluster yang menunjukkan tingkat keheterogenan antar kelompok



Gambar 13. Hasil visualisasi *scatter plot*

Kesimpulan

Penelitian ini bertujuan untuk melakukan klusterisasi data Top Rated Games on Steam menggunakan algoritma K-Means dengan metode elbow untuk menentukan jumlah cluster optimal. Proses klusterisasi dilakukan menggunakan perangkat lunak RapidMiner dengan tahapan yang melibatkan operator Select Attributes, Multiply, K-Means dengan 4 skenario jumlah cluster (2, 3, 5, dan 7), serta evaluasi menggunakan *Cluster Distance Performance*. Berdasarkan hasil evaluasi, jumlah cluster optimal yang diperoleh adalah tiga ($c = 3$). Ketiga cluster ini memiliki karakteristik sebagai berikut:

- a) Cluster 0 (Overwhelmingly Positive):
Game dengan ulasan positif sangat tinggi, namun jumlah ulasan negatif yang sangat rendah.
- b) Cluster 1 (Mostly Positive):
Game dengan ulasan positif yang dominan, tetapi disertai dengan jumlah ulasan negatif yang lebih tinggi dibandingkan cluster lainnya.
- c) Cluster 2 (Very Positive):
Game dengan ulasan positif dalam jumlah yang cukup signifikan dan ulasan negatif yang relatif lebih moderat.

Visualisasi hasil klusterisasi dalam bentuk scatter plot mempertegas perbedaan karakteristik antar cluster berdasarkan jumlah ulasan positif dan negatif. Penelitian ini berhasil mengelompokkan data ke dalam tiga cluster dengan baik, sehingga tujuan penelitian untuk mengidentifikasi pola ulasan game berdasarkan performa ulasan positif dan negatif dapat tercapai.

Hasil klusterisasi ini memberikan solusi dalam memahami tren ulasan game di Steam dan dapat dimanfaatkan oleh pengembang game sebagai referensi dalam menganalisis performa game mereka serta merancang strategi peningkatan kualitas produk yang lebih efektif.

Daftar Pustaka

- [1] p2dpt_uma, "Pentingnya Data dalam Era Digital: Menguak Potensi dan Tantangan." Accessed: Feb. 22, 2024. [Online]. Available: <https://p2dpt.uma.ac.id/2024/02/22/pentingnya-data-dalam-era-digital/>
- [2] C. A. Sugianto and K. K. Wanaziana, "Optimasi Algoritma K-Means Menggunakan Metode Elbow Pada Data Penerima Program Keluarga Harapan (PKH)," *INFORMATICS Digit. Expert*, vol. VOL. 6 NO., p. 65, 2024.
- [3] N. Rosyefa and Y. Rahmawati, "Implementasi RapidMiner Untuk Menentukan Siswa Unggulan Menggunakan Metode K-Means," p. 2, 2024.
- [4] Y. A. Adnan Burhan Hidayat Kiat and V. Rahmayanti, "Penerapan Metode K-Means Dengan Metode Elbow Untuk Segmentasi Pelanggan Menggunakan Model RFM (Recency, Frequency & Monetary)," vol. Vol. 2, No, p. 948, 2020.
- [5] Lorio, "RapidMiner," *BINUS Entrep. Cent.*, 2022, [Online]. Available: <https://binus.ac.id/entrepreneur/2022/12/08/rapidminer/>
- [6] Irfan_Satya_Aji, "Pengaruh Game Online Dan Pola Asuh Orangtua Terhadap Perilaku Agresif Siswa.," p. 7, 2020.
- [7] I. Bryan Arya Pangestu, "Analisa Kepuasan Pelanggan Pada Platform Game Steam Dengan Metode Importance Performance Analysis (IPA)," *Juminten*, vol. Vol. 03 No, pp. 49–60, 2022.

Jurnal.docx

ORIGINALITY REPORT

23%

SIMILARITY INDEX

20%

INTERNET SOURCES

9%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1	repositor.umm.ac.id Internet Source	6%
2	e-journal.upstegal.ac.id Internet Source	4%
3	publication.petra.ac.id Internet Source	2%
4	www.e-journal.unper.ac.id Internet Source	2%
5	juminten.upnjatim.ac.id Internet Source	2%
6	Rapel Sigit. "Penerapan Algoritma K-Means Clustering dalam Menganalisis Pola Peminjaman Buku di Perpustakaan", The Indonesian Journal of Computer Science, 2024 Publication	1%
7	publishing-widyagama.ac.id Internet Source	<1%
8	Lilik Harmaji, Akhmad Sufyan Asaury, Miki Wijana, Mohammad Erdda Habiby.	<1%

"Segmentasi Kinerja Akademik Mahasiswa Menggunakan Algoritma K-Means", Jurnal Accounting Information System (AIMS), 2024
Publication

9 Submitted to Universitas Sumatera Utara <1 %
Student Paper

10 Submitted to Universitas Muslim Indonesia <1 %
Student Paper

11 123dok.com <1 %
Internet Source

12 M. Padli, Amru Yasir, Buyung Solihin Hasugian. "PENERAPAN DATA MINING PADA MENU MAKANAN DAN MINUMAN KELOMPOK BERDASARKAN TINGKAT PENJUALAN MENGGUNAKAN K-MEANS", DEVICE : JOURNAL OF INFORMATION SYSTEM, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY, 2023 <1 %
Publication

13 repo.darmajaya.ac.id <1 %
Internet Source

14 Aqshol Al Fahrozi, Fitri Insani, Elvia Budianita, Iis Afrianty. "Implementasi Algoritma K-Means dalam Menentukan Clustering pada Penilaian Kepuasan Pelanggan di Badan Pelatihan Kesehatan Pekanbaru", Indonesian Journal of Innovation Multidisipliner Research, 2023 <1 %

15

download.garuda.kemdikbud.go.id

Internet Source

<1 %

16

www.researchgate.net

Internet Source

<1 %

17

Clarenza Dixie Rose, Bernadus Anggo Seno Aji, Farah Zakiyah Rahmanti. "Implementasi Aplikasi Web Pemilihan Kelas Berdasarkan Minat Menggunakan Algoritma K-Means Clustering", Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi), 2024

Publication

<1 %

18

Farid Raihan, Sutisna. "Analisis Pengaruh Media Sosial X Terhadap Sikap dan Perilaku Remaja Menggunakan K-Means Clustering", Jurnal Indonesia : Manajemen Informatika dan Komunikasi, 2024

Publication

<1 %

19

Fitri Khoirunnisa, Yulianto Rahmawati. "KOMPARASI 2 METODE CLUSTER DALAM PENGELOMPOKAN INTENSITAS BENCANA ALAM DI INDONESIA", Jurnal Informatika dan Teknik Elektro Terapan, 2024

Publication

<1 %

20

Idham Idham, Herliana Rosika, Yuliadi Yuliadi. "IMPLEMENTASI RAPIDMINER UNTUK CLESTERING DATA PENJUALAN PAKAIAN

<1 %

MENGGUNAKAN METODE K-MEANS", JUTECH : Journal Education and Technology, 2024

Publication

21

doaj.org

Internet Source

<1 %

22

jurnal.untan.ac.id

Internet Source

<1 %

23

ojs.unikom.ac.id

Internet Source

<1 %

24

repository.uksw.edu

Internet Source

<1 %

25

www.coursehero.com

Internet Source

<1 %

26

Heri Supriyanto, Mohammad Al Hafidz, Ari Cahaya Puspitaningrum, Rayhan Abdillah Putra Firmansyah, Rafi Zuhdi. "Klasterisasi Data Obat Farmasi Berdasarkan Jumlah Persediaan Dengan Menggunakan Metode K-Means", *Teknika*, 2024

Publication

<1 %

27

Martono Martono, Mohammad Syafrullah. "PENGKLASTERAN DAN SEGMENTASI KARAKTERISTIK DONATUR SEDEKAH DARING DENGAN TEKNIK PENAMBANGAN DATA", *INOVTEK Polbeng - Seri Informatika*, 2024

Publication

<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On

Jurnal.docx

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13
