



## *Comparison of K-Means and K-Medoids Algorithms for Grouping Video Game Sales Data in North America*

### **Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Penjualan Video Game di Amerika Utara**

**Aditya Rezky Pratama<sup>1</sup>, Bima Maulana<sup>2</sup>, Rahmad Didho Rianda<sup>3</sup>, Syaid El Hasyim<sup>4</sup>**

<sup>1,2,3,4</sup>Information System, State Islamic University of Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

E-Mail: <sup>1</sup>12050312773@student.uin-suska.ac.id, <sup>2</sup>12050313080@student.uin-suska.ac.id,  
<sup>3</sup>12050312613@student.uin-suska.ac.id, <sup>4</sup>12050312518@student.uin-suska.ac.id

*Makalah: Diterima 24 Juli 2023; Diperbaiki 09 Agustus 2023; Disetujui 27 September 2023*  
*Corresponding Author: Aditya Rezky Pratama*

#### **Abstract**

*Cluster analysis is a specific process of grouping a collection of objects into similar classes. Cluster is a group of objects that are similar to each other in the same cluster and different from objects in other clusters. Descriptive means that data mining is used to look for patterns that people can understand describing the properties of information. In addition, predictive data is used in building data models that are useful in making predictions. The K-Means Clustering algorithm is an important algorithm in the field of data mining and is easy to implement and run. In addition, the K-Means clustering method has a development variant, namely K-Medoids. The K-Medoids algorithm helps overcome the weaknesses of the K-Means algorithm. The amount of data that will be taken as a sample is 2000 data records. The data will be processed to obtain the average pattern of video game sales in North America. From the experiments conducted, it was found that K-means processing results with an average time of 1 second, whereas data processing with K-medoids takes 2 minutes 21 seconds on Rapidminer. The Davies-Bouldin Index (DBI) value on K-Means is lower than K-Medoids, namely 0.727 and 1.148 respectively, which means that the use of K-means is more appropriate for data processing of video game sales in North America.*

**Keyword:** Video game sales, Clustering, Data Mining, Davies-Bouldin Index, K-Means, K-Medoid

#### **Abstrak**

Analisis cluster merupakan proses spesifik pengelompokan kumpulan objek ke dalam kelas yang serupa. Kluster adalah sekumpulan objek yang mirip satu sama lain di dalam kluster yang sama dan berbeda dari objek di kluster lain. Deskriptif berarti data mining digunakan untuk menemukan pola yang bisa dipahami orang yang menggambarkan sifat-sifat informasi. Di samping itu data prediktif digunakan dalam membangun model data yang berguna dalam melakukan prediksi. Algoritma K-Means Clustering adalah algoritma yang penting dalam bidang data mining dan mudah untuk diimplementasikan dan dijalankan. Selain itu, metode pengelompokan K-Means memiliki varian pengembangan yaitu K-Medoids. Algoritma K-Medoids membantu mengatasi kelemahan dari algoritma K-Means. Banyak data yang akan diambil sebagai sampel adalah 2000 record data. Data tersebut akan diproses yang nantinya akan diperoleh pola rata-rata penjualan video game di Amerika Utara. Dari percobaan yang dilakukan, didapatkan hasil pengolahan K-means menggunakan Rapidminer dengan waktu yang diperlukan hanya rata-rata 1 detik sebaliknya pengolahan data dengan K-medoids memerlukan waktu 2 menit 21 detik. Nilai Davies-Bouldin Index (DBI) pada K-Means lebih rendah dibandingkan K-Medoids yaitu masing-masing 0.727 dan 1.148 yang berarti penggunaan K-means lebih tepat untuk pengolahan data penjualan video game di Amerika Utara.

**Kata Kunci:** Penjualan video game, Clustering, Data Mining, Davies-Bouldin Index, K-Means, K-Medoid

#### **1. PENDAHULUAN**

Video game sangat berkembang pesat dari tahun ke tahun. Semua masyarakat dari berbagai kalangan menjadikan video game sebagai hobi mereka untuk mendapatkan kesenangan tersendiri. Bahkan banyak orang yang ingin mensejahterakan hidupnya dengan cara melanjutkan karirnya sebagai atlet e-sport di bidang video game tertentu. Tentu hal ini memotivasi para developer untuk meraup keuntungan dengan cara membuat berbagai macam jenis dan genre game untuk menarik perhatian masyarakat.

Maka dari itu, para developer harus mengikuti perkembangan zaman dengan cara memanfaatkan data-data yang didapatkan dalam mengembangkan suatu video game. Hal tersebut dapat menjadi suatu informasi yang sangat penting sebagai acuan para developer dalam hal penjualan video game. Dengan informasi tersebut, developer dapat memprediksi perhitungan penjualan video game beserta mengantisipasi resiko yang akan dihadapi.

Dengan data tersebut, penulis dapat memperkirakan perhitungan penjualan video game yang didapatkan dengan cara mengelompokkan data penjualan video game menggunakan algoritma k-means dan k-medoid.

Analisis cluster merupakan proses spesifik pengelompokan kumpulan objek ke dalam kelas yang serupa. Kluster adalah sekumpulan objek yang mirip satu sama lain di dalam kluster yang sama dan berbeda dari objek di kluster lain. (Miller & Han, 2009). Analisis kluster atau cluster analysis merupakan sebuah fungsi dari data mining, dapat digunakan sebagai alat untuk mengelompokkan data. Secara umum, penggunaan data mining bisa dua bagian, yaitu deskriptif dan prediktif. Deskriptif berarti data mining digunakan mencari pola yang bisa dipahami orang menggambarkan sifat-sifat informasi. Di samping itu data prediktif digunakan dalam membangun model data yang berguna dalam melakukan prediksi. Bergantung pada fungsionalitas, tugas data mining dapat dibagi menjadi enam kelompok termasuk klasifikasi, Pengelompokan, regresi, deteksi anomali, analisis Asosiasi dan Ringkasan (Suyanto, 2017).

Dengan kuantitas data yang terkumpul terus meningkat dimana diperlukan alat dan teknik pengolahan data yang tepat agar pengumpulan data yang berlangsung bermanfaat untuk pengambilan keputusan di masa mendatang dalam suatu perusahaan atau organisasi. Dengan begitu adanya alat dan pengolahan yaitu data mining, dengan cepat mulai menjadi area penelitian yang penting dan diperlukan peningkatan penelitian ilmiah di bidang kecerdasan buatan dan statistik. [1][2].

Salah satu cara untuk mengelompokkan data selama data mining adalah dengan metode clustering. Pengelompokan ini bertujuan untuk mendapatkan pengelompokan dalam sekumpulan pola, titik, objek atau dokumen. Objek yang terdapat pada satu kelompok cluster yang sama mempunyai kesamaan antara satu kelompok dan perbedaan dengan objek pada kelompok cluster lainnya [4]. Algoritma K-Means Clustering adalah algoritma penting dalam data mining dan mudah diwujudkan dan dijalankan. Pada prakteknya, algoritma ini sering diterapkan karena relatif cepat dan adaptif [13]. Selain itu, metode pengelompokan K-Means memiliki varian pengembangan yaitu K-Medoids. Penggunaan K-Medoids yang bertujuan untuk mengurangi sensitivitas partisi yang dihasilkan terhadap nilai ekstrim yang terkandung dalam kumpulan data, tidak didasarkan pada rata-rata yang diamati dari setiap cluster. Algoritma K-Medoids membantu mencegah kekurangan dari algoritma K-Means. Algoritma K-Means sensitif terhadap outlier karena item dengan nilai besar dapat menyimpang secara signifikan dari distribusi data [14].

(Nanda, 2016) yang pada penelitiannya mencoba membandingkan antara K-Means dan K-Medoids dan memiliki saran penelitian kedepan yakni dengan mencoba memberikan hasil yang berbeda dengan menggunakan dataset yang berbeda.

Menurut (Dr. Aishwarya Batra, 2011) yang menganalisis dan memeriksa metode k-means dan k-medoid melalui pendekatan dasar disimpulkan bahwa algoritma K-mean lebih efisien untuk kumpulan data yang lebih kecil dibandingkan K-medoid yang kinerjanya lebih baik untuk kumpulan data besar.

Penelitian (Mediana Aryuni, 2018) sebelumnya dilakukan penelitian yg memperlihatkan lebih baiknya algoritma K-Means yakni pada jurnal berjudul segmentasi pelanggan pada Bank XYZ. Hasil penelitian ini memperlihatkan bahwa metode K-Means Clustering lebih baik metode K-Medoids berdasarkan jarak intra cluster (AWC). Sementara berdasarkan indeks Davies Bouldin, K-Means Clustering memiliki kinerja sedikit unggul daripada K-Medoids.

## **2. METODE DAN BAHAN**

### *2.1 Data Mining*

Data mining adalah proses menelusuri pola atau informasi yang menarik dalam data yang dipilih atau dikenal juga dengan istilah yang menggambarkan pencarian informasi dalam basis data dengan menerapkan metode atau teknik tertentu. Penambangan data atau data mining adalah proses menerapkan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengidentifikasi dan mengekstrak informasi berguna untuk informasi yang terkait dengan database berkapasitas tinggi. [5]. Salah satu metode yang

termasuk dalam data mining adalah clustering, yang berperan penting dalam pemeriksaan data, membuat prediksi dan mengatasi anomali data. Clustering adalah pengelompokan menerapkan teknik reiteratif dengan karakteristik kumpulan data yang identik [6].

## 2.2 Clustering

Pengelompokan data atau *clustering* adalah kegiatan yang memperhitungkan kemampuan penting untuk menemukan kesamaan dalam data dan membagi data yang sama ke dalam kelompok. *Clustering* dianggap sebagai bentuk terpenting dari pembelajaran tanpa pengawasan, di mana setiap masalah semacam ini terdiri dari menemukan struktur dalam kumpulan data yang tidak berlabel. *Clustering* membagi dataset menjadi beberapa kelompok, dimana kelompok tertentu lebih mirip dari kelompok lainnya [7][8].

Tipe data penting saat menggunakan algoritme pengelompokan untuk tujuan dan aplikasi tertentu. Eksperimen dengan algoritma yang berbeda pada data yang sama dapat dilakukan ketika menggunakan analisis cluster sebagai alat eksplorasi atau deskriptif yang tujuannya adalah untuk mendapatkan apa yang ditunjukkan oleh data. Secara umum, teknik pengelompokan dapat dibagi menjadi beberapa kategori, salah satunya adalah kategori teknik partisi. Teknik partisi ini dilandaskan pada penentuan pertama jumlah grup dan kemudian secara iteratif menugaskan ulang objek untuk menemukan grup pada titik tersebut. Salah satu algoritma yang paling populer untuk menerapkan metode partisi ini adalah algoritma K-Means dan algoritma KMedoids [8].

## 2.3 K-Means

K-means clustering adalah metode analisis cluster yang memiliki tujuan untuk membagi objek menjadi k cluster dan lalu mengamati dimana setiap objek pada cluster mendapatkan rata-rata berikutnya. Algoritma ini adalah salah satu solusi sederhana dan mudah dipelajari yang terkenal untuk masalah pengelompokan catatan. Algoritma K-Means merupakan algoritma evolusioner dimana operasinya memiliki arti yang sama dengan nama algoritmanya. Algoritma ini mengklasifikasikan observasi ke dalam k grup. di mana k adalah parameter masukan. Semua data kemudian ditempatkan ke dalam setiap klaster pengamatan sesuai dengan rata-rata kedekatan klaster pengamatan. Nilai rata-rata dalam kluster dihitung beberapa kali pada putaran pertama. Langkah-langkah pengelompokan K-means adalah sebagai berikut [8]:

1. Nilai k dipilih untuk pusat cluster awal secara acak.
2. Data yang ada pada dataset dibagi kedalam beberapa kelompok k cluster antara setiap titik dan pusat dari suatu cluster yang dihasilkan berdasarkan jarak Euclidean.

Berikut merupakan rumus perhitungan Euclidean Distance yang ditunjukkan pada persamaan (3.1) berikut [9]:

$$d(x,y) = \|x-y\|^2 = \dots\dots\dots(1)$$

dengan :

d= jarak data ke pusat cluster

x= data ke pada atribut ke

y= titik pusat ke

3. Setiap pusat cluster dihitung ulang berdasarkan nilai rata-rata pada cluster yang diperoleh.
4. Ulangi tahap 2 dan 3 hingga sesuai. Perulangan dapat diartikan secara berbeda tergantung pada implementasi, tetapi biasanya pada proses tahap 2 dan 3 dapat diulangi jika kelompok cluster masih ditemukan perubahan dan akan berhenti jika tidak ada perubahan antar material dalam cluster.

## 2.4 K-Medoids

Algoritma PAM (Partitioning Around Medoids) atau algoritma K-Medoids, adalah prosedur pemecahan yang diwakili oleh cluster yaitu medoid. Perbedaan antara prosedur pemecahan K-Medoids menggunakan prosedur pemecahan K-Means yaitu prosedur pemecahan K-Medoids memakai objek menjadi perwakilan (medoid) sentra cluster buat tiap cluster, ad interim prosedur pemecahan K-Means membutuhkan nilai rata-rata (mean) menjadi sentra cluster [17][10].

Berikut merupakan langkah-langkah penyelesaian algoritma K-Medoids [11]:

1. Inisialisasikan pusat cluster sebanyak jumlah cluster (k).
2. Setiap objek diposisikan ke cluster paling dekat menggunakan persamaan ukuran jarak Euclidian Distance dengan rumus persamaan (3.1).

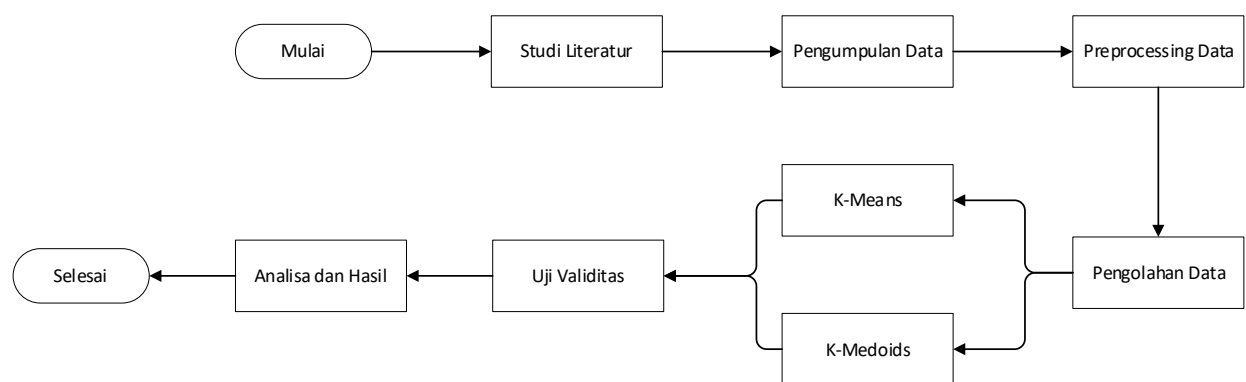
3. Objek yang ada pada masing-masing cluster dipilih secara acak yang kemudian ditetapkan sebagai calon medoid yang baru.
4. Hitung jarak setiap objek yang terdapat pada masing-masing cluster dengan kandidat medoid baru.
5. Total simpangan ( $S$ ) dihitung dengan cara nilai total jarak baru dikurangi dengan total jarak lama. Jika hasil  $S < 0$ , maka tukar objek dengan data cluster untuk membuat sekumpulan  $k$  objek baru sebagai medoid.
6. Ulangi tahap 3 sampai dengan 5 hingga tidak terjadi perubahan medoid, sehingga diperoleh cluster serta anggota cluster masing-masing.

### 2.5 Davies Bouldin Index (DBI)

Davies Bouldin Index (DBI) adalah matrik untuk menilai hasil algoritma klustering yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin bertepatan pada tahun 1979 [12]. Dengan menerapkan DBI suatu cluster dapat dianggap mempunyai skema clustering yang optimal yakni yang memiliki DBI minimal.

### 2.6 Metodologi Penelitian

Dilihat dari pendahuluan yang telah dipaparkan, beberapa tahap perlu dilalui dalam penyusunan penelitian ini, berikut tahapan kerangka kerja dalam penelitian ini:



**Gambar 1.** Metodologi Penelitian

Penelitian dimulai dengan melakukan studi literatur terkait penelitian yang telah dilakukan sebelumnya. Kemudian akan dilakukan pengumpulan data untuk kebutuhan penelitian, yang mana data-data yang telah didapat akan di preprocessing sesuai atribut-atribut yang dibutuhkan untuk penelitian. Selanjutnya dilakukan pengolahan data dengan cara menerapkan 2 algoritma yang berbeda terhadap data tersebut, yaitu algoritma k-means dan k-medoids. Penerapan algoritma tersebut akan menghasilkan tingkat akurasi yang akan diuji validitasnya, serta dilakukan analisa terhadap hasil pengolahan algoritma yang didapat. Maka dari itu, akan didapatkan kesimpulan mengenai pengolahan data dengan 2 algoritma yang berbeda

## 3. HASIL DAN PEMBAHASAN

### 3.1 Preprocessing Data

Data didapatkan dengan melakukan pencarian dari sebuah website kumpulan dataset yaitu kaggle. Data tersebut merupakan hasil analisa penjualan dari ribuan video game dari berbagai platform dan negara. Banyak data yang akan diambil sebagai sampel adalah 2000 record data. Data tersebut akan diproses yang nantinya akan diperoleh pola rata-rata penjualan video game di Amerika Utara. Dengan ini diharapkan bisa diterima oleh perusahaan sebagai informasi baru yang berguna dalam membuat sebuah kebijakan bisnis.

Dalam melakukan proses klasterisasi diperlukan sejumlah atribut. Dari 11 atribut yang ada pada data video game sale, diambil 5 atribut yang dipakai untuk proses perhitungan yaitu nama, genre, platform, publisher dan NA\_Sales. Atribut yang digunakan sebagai acuan untuk proses pengelompokan setelah dipilih tersebut genre, platform, publisher dan NA\_Sales, sedangkan atribut nama berguna sebagai kebutuhan analisa pengelompokan.

Cleaning data dilakukan agar noise dimimimalkan dikarenakan bisa mengubah perhitungan. Saat dilakukannya cleaning data akan dilakukan penghapusan data yang tidak cukup/lengkap, dengan begitu diperoleh sebanyak 2000 record data. Data hasil dari cleaning dapat dilihat pada tabel 1.

**Tabel 1. Cleaning Data**

| No | Genre        | Transformasi |
|----|--------------|--------------|
| 1  | Sports       | 0            |
| 2  | Platform     | 1            |
| 3  | Racing       | 2            |
| 4  | Role-Playing | 3            |
| 5  | Puzzle       | 4            |
| 6  | Misc         | 5            |
| 7  | Shooter      | 6            |
| 8  | Simulation   | 7            |
| 9  | Action       | 8            |
| 10 | Fighting     | 9            |
| 11 | Adventure    | 10           |
| 12 | Strategy     | 12           |

**Tabel 2. Transformasi Publisher**

| No.  | Nama                               | Platform | Genre        | Publisher            | NA_Sales |
|------|------------------------------------|----------|--------------|----------------------|----------|
| 1    | Wii Sports                         |          | Sports       | Nintendo             | 41.49    |
| 2    | Super Mario Bros.                  |          | Platform     | Nintendo             | 29.08    |
| 3    | Mario Kart Wii                     |          | Racing       | Nintendo             | 15.85    |
| 4    | Wii Sports Resort                  |          | Sports       | Nintendo             | 15.75    |
| 5    | Pokemon Red/Pokemon Blue           |          | Role-Playing | Nintendo             | 11.27    |
| .... | ....                               |          | ....         | ....                 | ....     |
| 24   | Super Mario Bros. 3                |          | Platform     | Nintendo             | 9.54     |
| 25   | Grand Theft Auto V                 |          | Action       | Take-Two Interactive | 9.63     |
| .... | ....                               |          | ....         | ....                 | ....     |
| 2000 | Dynasty Warriors 3: Xtreme Legends |          | Action       | Tecmo Koei           | 0.13     |

Setelah dilakukannya Cleaning data, Tabel 2. Transformasi Publisher, berikutnya dilakukan transformasi data. Dalam memudahkan perhitungan nantinya, data perlu dilakukan transformasi menjadi bentuk simbol. hasil dari perubahan data menjadi bentuk simbol bisa dilihat pada tabel 2 sampai dengan tabel 4.

**Tabel 3. Transformasi Genre**

| No. | Publisher                   | Transformasi |
|-----|-----------------------------|--------------|
| 1   | Nintendo                    | 0            |
| 2   | Microsoft Game Studios      | 1            |
| 3   | Take-Two Interactive        | 2            |
| 4   | Sony Computer Entertainment | 3            |
| 5   | Activision                  | 4            |
| 6   | Ubisoft                     | 5            |
| 7   | Bethesda Softworks          | 6            |
| 8   | Electronic Arts             | 7            |
| 9   | Sega                        | 8            |
| 10  | SquareSoft                  | 9            |
| 11  | Atari                       | 10           |
| 12  | 505 Games                   | 11           |
| 13  | Capcom                      | 12           |
| 14  | GT Interactive              | 13           |

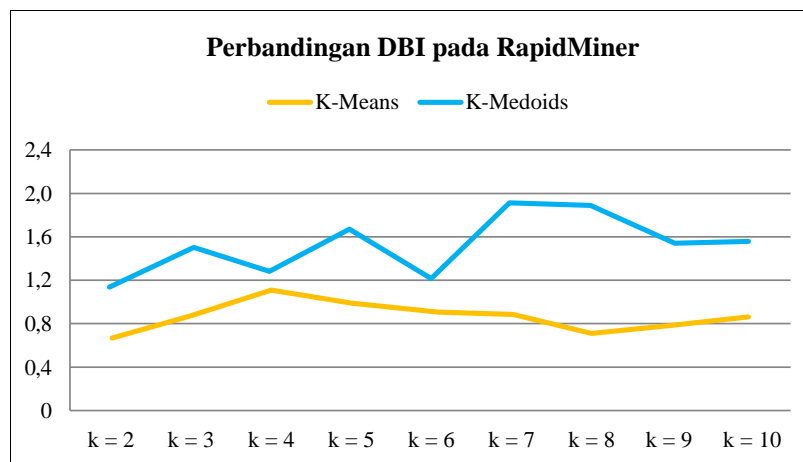
|      |                              |      |
|------|------------------------------|------|
| 15   | Konami                       | 14   |
| 16   | Sony<br>Entertainment Europe | 15   |
| .... | ....                         | .... |
| 96   | GSV                          | 95   |
| 97   | Microprose                   | 96   |
| 98   | Play It                      | 97   |

Berdasarkan atribut pada tabel 2 sampai dengan 3 maka dapat diperoleh transformasi secara keseluruhan, yang ditunjukkan pada tabel 4:

**Tabel 4.** Normalisasi

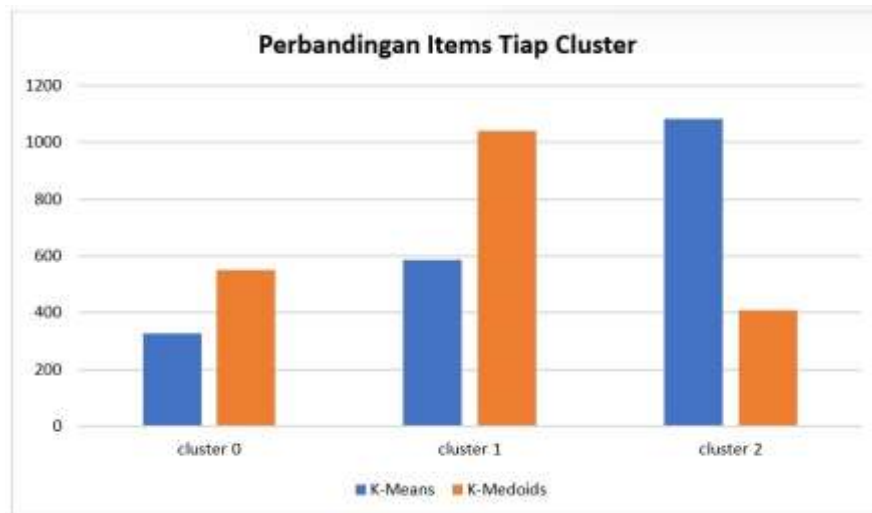
| No.  | Publisher | Platform | Genre  | NA_Sales |
|------|-----------|----------|--------|----------|
| 1    | -1.448    | -1.446   | -0.780 | 20.108   |
| 2    | -1.138    | -1.250   | -0.780 | 13.888   |
| 3    | -0.829    | -1.446   | -0.780 | 7.256    |
| 4    | -1.448    | -1.446   | -0.780 | 7.206    |
| 5    | -0.519    | -1.054   | -0.780 | 4.961    |
| ...  | ...       | ...      | ...    | ...      |
| 24   | -0.642    | -0.662   | 1.031  | 4.139    |
| 25   | -0.642    | -0.270   | 1.031  | 3.527    |
| ...  | ...       | ...      | ...    | ...      |
| 2000 | 1.139     | -0.270   | 1.031  | -0.533   |

Sesudah dilakukan preprocessing data, data berikutnya diolah dengan RapidMiner agar diperoleh nilai DaviesBouldin Index (DBI) yang menjadi acuan pengelompokan cluster terbaik. Cluster  $k=2$  sampai dengan  $k=10$  menjadi objek pengujian. Berikut perolehan dari pengolahan data yang diperlihatkan pada gambar 2:



**Gambar 2.** Metodologi Penelitian

Dilihat dari perbandingan nilai Davies-Bouldin Index diatas, menyatakan bahwa yang menjadi cluster terbaik adalah  $k=2$ . Pembagian tiap items pada masing-masing pengelompokan  $k=2$  dapat diperlihatkan pada gambar 3.



Gambar 3. Metodologi Penelitian

Dari percobaan menggunakan Rapidminer, didapatkan hasil pengolahan K-means dengan waktu yang diperlukan hanya rata-rata 1 detik sebaliknya pengolahan data dengan K-medoids membutuhkan waktu 2 menit 21 detik pada Rapidminer. Nilai Davies-Bouldin Index (DBI) pada K-Means lebih rendah dibandingkan K-Medoids yaitu masing-masing 0.727 dan 1.148 yang berarti penggunaan K-means lebih tepat untuk pengolahan data penjualan video game di Amerika Utara.

#### 4. KESIMPULAN

Pengelompokkan data yang ditunjukkan dari perbandingan antara algoritma clustering, K-Medoids dan K-Means tidak terlalu signifikan. Rata-rata waktu yang diperlukan yakni pada K-Means hanya 1 detik sedangkan pada K-Medoids memerlukan 2 menit 21 detik dengan artian semakin tinggi iterasi dari grup/kelompok yang ditentukan, maka pemrosesan data akan semakin lambat. Kemudian nilai Davies-Bouldin dibandingkan Index diperoleh nilai  $k=3$  pada K-Means lebih rendah dibandingkan KMedoids. Lalu pada  $K=10$  juga menunjukkan pengolahan data berlanjut dengan nilai DBI yang ditunjukkan K-Means lebih rendah K-Medoids. Pola pengelompokkan masing-masing cluster didapatkan dari NA\_Sales (Penjualan video game pada Amerika Utara).

#### References

- [1] Arora P, D. D. (2016). Analysis of K-Means and K-Medoids Algorithm For Big Data. *International Conference on Information Security & Privacy (ICISP2015)*, (pp. 507-512).
- [2] Aryani, M., & dkk. (2018). Customer Segmentation in XYZ Bank using K-Means and K-Medoids Clustering. *International Conference on Information Management and Technology (ICIMTech)*, (pp. 412-416). Jakarta.
- [3] Balabantaray RC, S. C. (2013). Document clustering using K-Means and K-Medoids. *International Journal of Knowledge Based Computer System*, 7-13.
- [4] Batra, D. A. (2011). *Analysis and Approach: K-Means and K-Medoids Data Mining Algorithms*. India.
- [5] Han J, K. M. (2006). *Data Mining: Concept and Techniques*. Morgan Kauffman.
- [6] Han J, K. M. (2012). *Data Mining : Concepts and techniques*. USA: Elsevier Inc.
- [7] Han, M. &. (2009). *Geographic Data Mining and Knowledge Discovery*. Boca Raton.
- [8] Herviany, M ., & dkk. (2021). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokkan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat, *Indonesian Journal of Machine Learning and Computer Science (MALCOM)*, 34-40
- [9] Kaur, N. K. (2014). K-Medoids Clustering Algorithm.
- [10] Kurmiati, D., & dkk (2021) Klasterisasi Daerah Rawan Gempa Bumi di Indonesia Menggunakan Algoritma K-Medoids, *Indonesian Journal of Machine Learning and Computer Science (MALCOM)*, 47-57
- [11] Luchia, T. N., & dkk (2022) Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia, *Indonesian Journal of Machine Learning and Computer Science (MALCOM)*, 35-41
- [12] MT., U. H. (2017). Students Academic Performance Using Partitioning Clustering Algorithms. *International Journal of Advanced Research in Computer Science*, 640-644.

- [13] Nanda, A. (2016). Comparative Study between Parallel K-Means and Parallel K-Medoids using Message Passing Interface (MPI). *International Journal on Information and Communication of Technology*, 27-36.
- Okfalisa, M. G. (2017). Comparative Analysis Of K-Nearest Neighbor. *Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 294-298.
- [14] Singh K, M. D. (2011). Evolving Limitations in K-Means Algorithm in Data Mining and Their Removal. *International Journal of Computational Engineering & Management (IJCEM)*, 105-109.
- [15] Suyanto. (2017). *Data Mining Untuk Klasifikasi dan Klusterisasi Data*. Bandung: INFORMATIKA.
- [16] TS, M. (2011). Comparison Between K-Means and K-Medoids Clustering Algorithms. *International Journal of Advanced Computing (IJAC)*, 65-69.
- [17] WA, T. (2015). Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk. *Jurnal Teknik Industri, Mesin, Elektro, dan Ilmu Komputer*, 183-188.
- [18] Wu, X. &. (2009). *The Top Ten Algorithms in Data Mining*. London: CRC Press.