

Analisa Clustering untuk Mengelompokan Data Penayangan Film Bioskop Menggunakan Algoritma K-Means

Moh. Nurdayat¹, Nana Suarna², Yudhistira Arie Wijaya³

^{1,3}Teknik Informatika, STMIK IKMI Cirebon, Indonesia

³Sistem Informasi, STMIK IKMI Cirebon, Indonesia

mohammadnurdayat43@gmail.com

Info Artikel

Sejarah artikel:

Diterima Mei 2023

Direvisi Juni 2023

Disetujui Juli 2023

Diterbitkan Juli 2023

ABSTRACT

The purpose of this study is one of the analyzes to obtain film screening data, the approach used in this study is the K-means algorithm using the parameter measure type Numerical Measure with Numerical Measure Euclidean Distance to get the best Davies Bouldin Index (DBI), with the intention of getting helps grouping datasets of film screenings at the Ramayana Cirebon XXI Cinema. Results from the evaluation of the Davies Bouldin Index (DBI) obtained is (K-2) with a Davies Bouldin Index (DBI) value of 0.864, because the value obtained is the smaller the Davies Bouldin Index (DBI) value, it shows the optimum performance of the resulting cluster.

Keywords : Clustering; Davies Bouldin Index; K-Means.

ABSTRAK

Tujuan penelitian ini yaitu menganalisa clustering untuk memperoleh data penayangan film, pendekatan yang digunakan dalam penelitian ini adalah algoritma *K-means* dengan menggunakan parameter measur type jenis *Numerical Measure* dengan *Numerical Measure Euclidean Distance* untuk mendapatkan *Davies Bouldin Index (DBI)* terbaik, dengan maksud dapat membantu pengelompokan dataset penayangan film di Bioskop Ramayana Cirebon XXI. Hasil dari evaluasi *Davies Bouldin Index (DBI)* yang diperoleh yaitu (K-2) dengan nilai *Davies Bouldin Index (DBI)* 0,864, karena nilai yang didapat semakin kecil nilai *Davies Bouldin Index (DBI)* maka menunjukkan performa optimum cluster yang dihasilkan.

Kata Kunci : *Clustering; Davies Bouldin Index; K-Means.*

PENDAHULUAN

Industri film semakin hari semakin banyak salah satunya Indonesia sebagai salah satu pangsa pasar utama di dunia. Penelitian ini penting untuk diteliti dimaksudkan untuk mengetahui kelompok data atau cluster tiap penayangan film, dengan melakukan teknik *Data Mining* untuk melihat hasil cluster, *Data Mining* adalah proses menemukan pola dan tren yang berguna dalam dataset yang besar [1]. Metode algoritma *K-means* digunakan dengan menentukan nilai *Davies Bouldin Index (DBI)* terbaik, pengelompokan menggunakan Algoritma *K-means* dimulai dengan titik K yang dilakukan sebagai pusat massa dari K cluster potensial, titik pertama ini adalah posisi yang diambil secara acak, selanjutnya centroid dari cluster dihitung ulang dengan meratakan semua dataset dari satu cluster sampai centroid tidak lagi bergerak atau langkah optimalisasi maksimal tercapai [3]. *Davies Bouldin Index*

(DBI) merupakan alat ukur untuk mengevaluasi algoritma pengelompokan sebagai skema evaluasi dimana validasi seberapa baik pengelompokan telah dilakukan menggunakan jumlah dan fitur pada dataset, *Davies Bouldin Index* (DBI) bekerja dengan menghitung nilai rata-rata setiap item dalam kumpulan data, jumlah klaster yang dipilih adalah jumlah klaster yang memiliki nilai *Davies Bouldin Index* (DBI) terkecil dan menunjukkan bahwa itulah nilai *Davies Bouldin Index* (DBI) terbaik [4]. Dengan metode *K-means* menggunakan parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance* ini diharapkan dapat menghasilkan kelompok atau cluster dengan nilai *Davies Bouldin Index* (DBI) yang terbaik [5].

Penelitian terdahulu menurut Mulyadi dan Iriene Putri yang berjudul "Klasterisasi Menggunakan Metode Algoritma K-Means dalam Meningkatkan Penjualan Tupperware" Tujuan dari penelitian merupakan sarana untuk membantu Toko dalam meningkatkan penjualan tupperware dengan menghasilkan produk yang sangat laris dan tidak sama sekali laris, dari penelitian ini di dapatkan bahwa metode algoritma *K-means* layak dan efektif untuk pengelompokan 3 cluster, Cluster 1 (C1) Sangat Laris, Cluster 2 (C2) Laris, dan Cluster 3 (C3) Tidak Laris [6]. Metode *K-means* dapat diterapkan untuk mengelompokkan data terbukti dengan hasil penelitian dari Nofiar, Andri Defit, dan Sarjon Sumijan, dengan hasil penelitian yang memberikan cluster 0 kategori baik berjumlah 12 data, cluster 1 kategori sangat baik berjumlah 13 data, dan cluster 2 kategori kurang baik berjumlah 6 data, diperoleh dari hasil metode *K-means Clustering* dapat digunakan untuk proses pengolahan data menggunakan konsep *Data Mining* dalam mengelompokkan data sesuai criteria [7].

Usulan penelitian ini merupakan optimalisasi parameter pada algoritma *K-means*, sumber data yang digunakan dalam penelitian ini adalah dataset penayangan film di Bioskop, terhitung dari bulan Oktober 2022 sampai Januari 2023, kontribusi atau novelty yang akan dilakukan pada penelitian ini adalah dengan menggunakan metode *K-means*, langkah-langkah yang akan dilakukan dengan menentukan nilai ($K = 2, 3, 4, 5$) dengan berdasarkan parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance* sampai menghasilkan nilai *Davies Bouldin Index* (DBI) terbaik.

Kajian penelitian ini menjelaskan sumber data terbuka yang populer digunakan alat bantu *Rapid Miner*, pengelompok menggunakan algoritma *K-means* dalam pengelompokan dataset penayangan film Bioskop, pengelompokan data dapat dijadikan sebagai bagian dari penambahan data sumber terbuka digunakan dalam pengolahan data skala besar. [8].

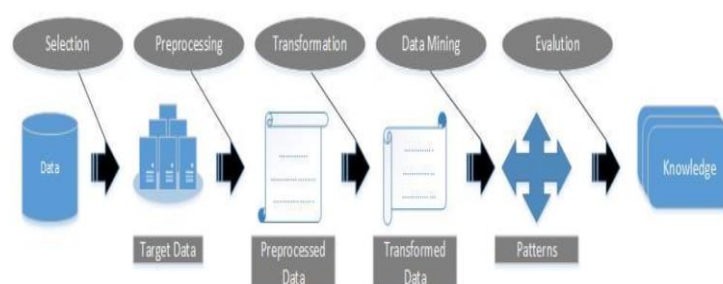
Tujuan utama dari penelitian ini merupakan salah satu analisa untuk memperoleh data penayangan film, pendekatan yang digunakan dalam penelitian ini adalah algoritma *K-means* dengan maksud dapat membantu pengelompokan dataset penayangan film untuk mendapatkan nilai *Davies Bouldin Index* (DBI) yang terbaik di Bioskop Ramayana Cirebon XXI, eksperimen akan dilakukan dengan menggunakan parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance*, dengan mengcluster dari

beberapa film yang ditayangkan, kemudian data referensi digunakan untuk menjadi estimasi asli [9].

Pengelompokan data penayangan film sangat tepat dengan sistem *Clustering* sehingga lebih mudah memahami genre film yang diminati para penonton, bentuk media massa paling populer di negara ini adalah film, berbagai jenis genre film yang dirilis baru-baru ini telah memberikan rasa penasaran pada penonton untuk menonton film yang ingin mereka tonton [10].

METODE

Pada penelitian ini, menggunakan metode *Knowledge Discovery In Database* (KDD) Berikut merupakan tahapan - tahapan dan proses yang ada pada *Knowledge Discovery In Database*.



Sumber : H.Syahputra [11]

Gambar 1. Tahapan *Knowledge Discovery In Database*

Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery In Database* dimulai, proses seleksi dilakukan pada data penayangan film di bioskop Ramayana Cirebon XXI, terhitung dari bulan Oktober 2022 sampai Januari 2023 berjumlah 9 atribut dan berisi 492 data, untuk pengolahan data tersebut, maka pada tahap selection ini pengolahan data dilakukan menggunakan *Microsoft Excel 2007* dengan format *xlsx* untuk melihat data yang diperoleh kemudian import data kedalam aplikasi *Rapid Miner* serta melakukan pembacaan data menggunakan *Read Excel* dan *Set Role*, kemudian memilih atribut yang digunakan menggunakan *Select Attribute*.

Pre-processing

Preprocessing/cleaning bertujuan untuk membuang duplikasi data, memeriksa data yang tidak konsisten dan memperbaiki kesalahan pada dataset, proses ini merupakan tahapan penting karena jika tidak dilakukan dapat menghambat hasil dari proses *Data Mining*. tetapi hasil menunjukkan pada dataset ini tidak memiliki *Missing Value* atau kesalahan pada dataset.

Transformation

Transformasi dilakukan dengan menggunakan empat operator, pertama data yang berjenis nominal dirubah terlebih dahulu kedalam jenis angka, pada

penelitian ini menggunakan 9 atribut yang diolah menggunakan aplikasi *Rapid Miner* atribut jenis, Tanggal, Genre Film, HTM, Show 1, Show 2, Show 3, Theater, FP (Freepass) dan Uang, diantara dua atribut tersebut akan dilakukan transformasi data jenis *Polynomial* menjadi jenis *Numerical* dengan menggunakan tools *Nominal To Numerical* yaitu atribut "Genre Film", kemudian merubah atribut "Tanggal" dari jenis *Date* menjadi jenis *Numerical* dengan menggunakan operator *Date To Numerical*, selanjutnya melakukan langkah normalisasi kepada atribut yang dipilih menggunakan operator *Normalize*, secara proses menghaluskan dataset dengan menggunakan operator *Exponential Smoothing*, dengan tujuan untuk menyiapkan dataset, dengan cara mengubah data yang sudah dipilih dan dikonversi menjadi bentuk yang cocok untuk proses *Data Mining*.

Data Mining

Data mining merupakan proses untuk mencari pola atau informasi yang ada didalam dataset yang sudah terpilih pada saat proses transformasi data, pengujian pada dataset yang telah ditentukan memerlukan suatu metode yang digunakan yaitu *K-means* dari software *Rapid Miner* yang digunakan sesuai dengan tujuan penelitian, langkah-langkah yang akan dilakukan dengan menentukan nilai ($K = 2, 3, 4, 5$) dengan berdasarkan parameter measure type *Numerical Measures* dengan *Numerical Measures Euclidean Distance* dengan dibantu operator *Cluster Distance Performance* hingga menghasilkan *Davies Bouldin Index (DBI)* terbaik.

Interpretation / Evaluation

Pola Informasi yang diperoleh dari proses *Data Mining* dilakukan proses evaluasi terhadap data yang sudah didapat sesuai dengan prosedur yang telah ditentukan, tahap ini juga menggunakan aplikasi *Rapid Miner* untuk mengukur kemampuan metode algoritma *K-means* dengan uji *performace* hingga menghasilkan nilai *Davies Bouldin Index* terbaik.

HASIL DAN PEMBAHASAN

Data Selection

Data selection merupakan tahap pertama dalam proses *Knowledge DiscoveyIn Database (KDD)*, dimana dari sekumpulan data yang diperoleh dilakukan seleksi data terlebih dahulu, data dalam bentuk file *Microsoft Excel* dengan format *xlsx* yang diperoleh kemudian import data kedalam aplikasi *Rapid Miner* untuk melakukan pembacaan data menggunakan operator *Read Excel* seperti pada gambar 1 dibawah ini.



Gambar 1. Operator Read Excel

Data selection seperti pada tabel 1 berikut :

Tabel 1. Statistik Data Read Excel

No.	Uraian	Keterangan
1.	Record	492
2.	SpecialAttribute	0
3.	Reguler Attribute	10
4.	Attributes :	
	NO	Interger, missing 0
	Tanggal	Date, missing 0
	Genre Film	Polynominal,missing 0
	HTM	Interger, missing 0
	Show 1	Interger, missing 0
	Show 2	Interger, missing 0
	Show 3	Interger, missing 0
	Theater	Interger, missing 0
	FP (Freepass)	Interger, missing 0
	Uang	Interger, missing 0

Pre-processing

Langkah selanjutnya adalah melakukan *pre-processing* pada data yang memiliki *Missing Value*, ketika sebuah atribut dalam kumpulan data yang sudah dipilih tidak memiliki nilai atau kosong maka itu diindikasikan sebagai *Missing Value*, tetapi hasil menunjukkan pada dataset ini tidak memiliki *Missing Value* berdasarkan hasil dari statistik dataset pada tabel 2, diketahui bahwa atribut memiliki nilai 0 pada tiap kolom missing, maka pada tahap preprocessing tidak perlu dilakukan.

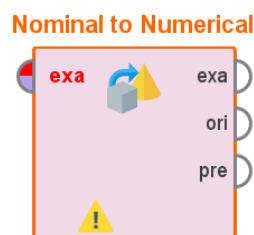
Tabel 2. Statistik Data Tanpa Ada MissingVlue

No.	Uraian	Keterangan
1.	Record	492
2.	SpecialAttribute	1
3.	Reguler Attribute	9
4.	Attributes :	
	NO (Id)	Interger, missing 0
	Tanggal	Date, missing 0
	Judul Film	Polynominal,missing 0
	HTM	Interger, missing 0
	Show 1	Interger, missing 0
	Show 2	Interger, missing 0
	Show 3	Interger, missing 0
	Theater	Interger, missing 0
	FP (Freepass)	Interger, missing 0
	Uang	Interger, missing 0

Transformation

Tahap transformasi dilakukan dengan menggunakan empat operator dengan tujuan untuk menyiapkan dataset, dengan cara mengubah data yang sudah dipilih dan dikonversi menjadi bentuk yang cocok untuk proses *Data*

Mining, proses transformasi menggunakan operator *Nominal To Numerical*, *Date To Numerical*, *Normalize* dan *Exponential Smoothing*.



Gambar 2. Operator *Nominal to Numerical*

Operator *Nominal To Numerical* membantu proses dataset untuk dapat di proses pada *Data Mining* dengan cara merubah data *Polynomial* menjadi *Numerical*, tahap ini dilakukan bertujuan untuk memudahkan proses *Clustering*, salah satu atribut yang dirubah yaitu "Genre Film", karena atribut genre film berjenis nominal yang nantinya tidak dapat diproses menggunakan operator *K-means* dengan parameter *measure type Numerical Measure*.

Tabel 3. Parameter Operator *Nominal to Numerical*

No.	Parameter	Isi
1.	Attribute Filter Type	Subset
2.	Attributes	Genre Film

Hasil dari pembacaan operator *Nominal to Numerical* didapat informasi yang dapat dilihat pada tabel 4.

Tabel 4. Statistik Data *Nominal to Numerical*

No.	Uraian	Keterangan
1.	Record	492
2.	SpecialAttribute	1
3.	Reguler Attribute	24
4.	Attributes :	
	NO (Id)	Interger, missing 0
	Genre Film	Numerical, missing 0
	Tanggal	Date, missing 0
	HTM	Interger, missing 0
	Show 1	Interger, missing 0
	Show 2	Interger, missing 0
	Show 3	Interger, missing 0
	Theater	Interger, missing 0
	FP (Freepass)	Interger, missing 0
	Uang	Interger, missing 0

Data Mining

Data mining merupakan proses untuk mencari pola atau informasi yang ada didalam dataset yang sudah terpilih pada saat proses transformasi, penelitian ini menggunakan *K-means* dengan mengcluster data penayangan film di Bioskop, operator *K-means* dapat dilihat pada Gambar 3 dibawah ini.



Gambar 3. Operator K-Means

Algoritma *K-means* digunakan untuk menentukan cluster dari data yang sudah dipilih, dengan melakukan terlebih dahulu pemilihan jumlah cluster yang berbeda yaitu ($K = 2, 3, 4, 5$), jumlah cluster yang digunakan dalam upaya bereksperimen terhadap parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance*.

Tabel 4. Parameter K-Means

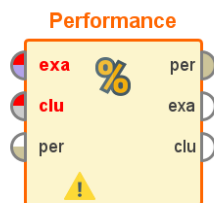
No.	Parameter	Isi
1.	K	2, 3, 4, 5
2.	Max Runs	10
3.	MeasureTypes	Numerical Measure
4.	NumericalMeasure	Euclidean Distance
5.	Max Optimization	100

Hasil dari pembacaan operator *K-means* didapat informasi yang dapat dilihat pada tabel 5.

Tabel 5. Statistik Data K-means

No.	Uraian	Keterangan
1.	Record	492
2.	SpecialAttribute	2
3.	Reguler Attribute	9
4.	Attributes :	
	NO, (Id)	Integer, missing 0
	Genre Film	Real, missing 0
	HTM	Real, missing 0
	Show 1	Real, missing 0
	Show 2	Real, missing 0
	Show 3	Real, missing 0
	Theater	Real, missing 0
	FP (Freepass)	Real, missing 0
	Uang	Real, missing 0

Menambahkan operator *Cluster Distance Performance* pada proses *Clustering* untuk mengevaluasi kinerja proses *Clustering* dengan metode *Davies Bouldin Index (DBI)*, operator *Cluster Distance Performance* dapat dilihat pada Gambar 4. *Cluster Distance Performance* digunakan untuk menentukan hasil dari cluster set ($K = 2, 3, 4, 5$) dengan berdasarkan parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance* pada *K-means*, sampai data yang dikelompokkan mendapatkan nilai terkecil berdasarkan *Davies Bouldin Index (DBI)*.



Gambar 4. Cluster Distance Performance

Tabel 6. Parameter Cluster Distance Performance

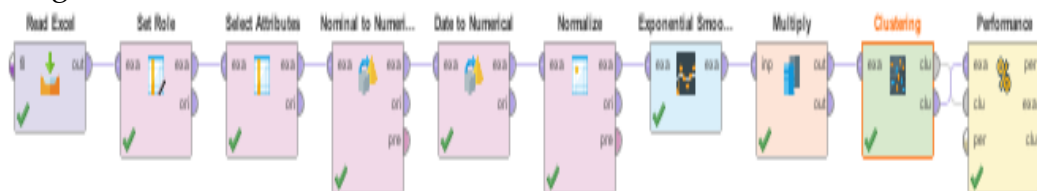
No.	Parameter	Isi
1.	Main Criterion	DaviesBouldin Index

Hasil dari pembacaan operator Cluster Distance Performance dapat dilihat pada tabel 6.

Tabel 7. Statistik Data Cluster Distance Performance

No.	Uraian	Keterangan
1.	Record	492
2.	SpecialAttribute	2
3.	Reguler Attribute	9
4.	Attributes :	
	NO, (Id)	Interger, missing 0
	Cluster	Nominal, missing 0
	Genre Film	Real, missing 0
	HTM	Real, missing 0
	Show 1	Real, missing 0
	Show 2	Real, missing 0
	Show 3	Real, missing 0
	Theater	Real, missing 0
	FP (Freepass)	Real, missing 0
	Uang	Real, missing 0

Dataset yang telah ditentukan memerlukan penggunaan satu model dalam metode K-means, proses model K-means dengan operator Cluster Distance Performance bertujuan untuk mendapatkan nilai Davies Bouldin Index (DBI) dari masing-masing cluster set (K = 2, 3, 4, 5) yang sudah ditentukan, model K-means seperti gambar 5 dibawah ini.



Gambar 5. Proses Model K-means

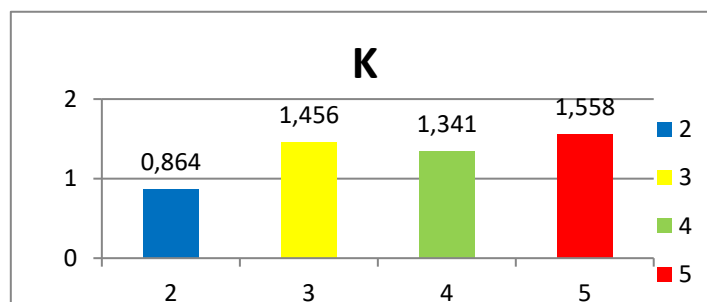
Pengujian dengan menggunakan Data Mining membantu untuk proses Clustering menjadi lebih mudah dan efektif, proses ini merupakan data yang dikelompokkan berdasarkan kemiripan karakteristik dari setiap data.

Interpretation/Evaluation

Evaluasi cluster dilakukan untuk menentukan jumlah cluster yang berhasil menghasilkan nilai *Davies Bouldin Index* terkecil, atau jumlah cluster dengan tingkat kemiripan data terdekat, rekapitulasi *Davies Bouldin Index* (DBI) yang diperoleh dengan menggunakan cluster set (K = 2, 3, 4, 5) parameter measure type *Numerical Measure* dengan *Numerical Measure Euclidean Distance* sebagai hasil pengujian seperti tabel 8 berikut.

Tabel 8. Hasil Evaluasi Cluster

MeasureType	Cluster Set	DBI	Jumlah Anggota Cluster
Numerical Measure	2	0,864	C_0 : 333 Item C_1 : 159 Item Jumlah Item : 492
	3	1.456	C_0 : 143 Item C_1 : 237 Item C_2 : 112 Item Jumlah Item : 492
	4	1.341	C_0 : 192 Item C_1 : 88 Item C_2 : 65 Item C_3 : 147 Item Jumlah Item : 492
	5	1.558	C_0 : 62 Item C_1 : 71 Item C_2 : 100 Item C_3 : 183 Item C_4 : 76 Item Jumlah Item : 492



Gambar 6. Grafik (K = 2, 3, 4, 5)

Berdasarkan hasil rekapitulasi *Davies Bouldin Index* (DBI) yang ditunjukkan pada tabel 8 dan grafik 6, dengan hasil *Davies Bouldin Index* (DBI) yang diperoleh yaitu (K-2) dengan hasil 0,864, (K_3) dengan hasil 1,456, (K-4) dengan hasil 1,341 dan (K-5) hasil 1,558, maka cluster set yang akan digunakan adalah cluster set (K-2) karena memiliki nilai *Davies Bouldin Index* (DBI) terkecil yaitu 0,864, karena nilai yang didapat semakin kecil nilai *Davies Bouldin Index* (DBI) maka menunjukkan performa optimum cluster yang dihasilkan.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -13.679
Avg. within centroid distance_cluster_0: -13.351
Avg. within centroid distance_cluster_1: -14.366
Davies Bouldin: -0.864
```

Gambar 7. Hasil DaviesBouldin Index (DBI) (K-2)

Hasil uji performance dengan evaluasi metode *Davies Bouldin Index* (DBI) dalam kinerja pengelompokan cluster dalam menguji data penayangan film, berdasarkan hasil *Davies Bouldin Index* (DBI) yang diperoleh yaitu (K-2) dengan nilai *Davies Bouldin Index* (DBI) 0,864, terdapat satu hasil cluster model yang diperoleh dan ditunjukkan pada Gambar 8.

Cluster Model

```
Cluster 0: 333 items
Cluster 1: 159 items
Total number of items: 492
```

Gambar 8. Hasil Cluster Set (K-2)

Hasil cluster set (K-2) dengan menunjukkan cluster_0 memiliki 333 data, cluster_1 memiliki 159 data, dan jumlah keseluruhan 492 data.

Tabel 9. Hasil Centroid Akhir (K-2)

Attribute	Cluster_0	Cluster_1
Genre Film	4.226	4.197
HTM	1.184	9.029
Show 1	2.105	3.872
Show 2	2.390	3.987
Show 3	1.773	2.179
Theater	5.029	4.871
FP (Freepass)	0.412	0.218
Uang	0.798	1.549

PENUTUP

Berdasarkan hasil dari penerapan data *mining* pada data penayangan film di bioskop Ramayana Cirebon XXI menggunakan algoritma *K-Means*, maka dapat disimpulkan sebagai berikut :

1. Pengelompokan data penayangan film di bioskop Ramayana Cirebon XXI menggunakan algoritma *K-means*, menghasilkan cluster set (K-2) dengan 2 cluster, cluster_1 didominasi oleh genre film Horor sebesar 58%, Fantasi 33%, Drama 8% dan Action 1% dan cluster_0 dari Genre Film Horor sebesar 57%, Fantasi 31%, Drama 9% dan Action 3%, adapun nilai *Davies Bouldin Index* (DBI) yang didapatkan 0,864.
2. Algoritma *K-means* dapat diimplementasikan pada proses *Clustering* menggunakan software *Rapid Miner*, pengujian cluster dengan menentukan cluster set (K = 2, 3, 4, 5) pada dataset penayangan film dengan proses *Clustering* berdasarkan parameter measure type *Numerical Measure* dengan

Numerical Measure Euclidean Distance, mendapatkan nilai *Davies Bouldin Index* (DBI) yang diperoleh pada cluster set (K-2) menghasilkan 2 cluster dengan nilai *Davies Bouldin Index* (DBI) 0,864.

DAFTAR PUSTAKA

- [1] H. Pandiangan, "Penerapan Data Mining Dalam Clustering Produksi Daging Sapi Di Indonesia Menggunakan Algoritma K-Means," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 1, no. 2, pp. 37-44, 2019, [Online]. Available: <https://doi.org/10.47709/cnipc.v1i2.239>
- [2] H. Pandiangan, "Penerapan Data Mining Dalam Clustering Produksi Daging Sapi Di Indonesia Menggunakan Algoritma K-Means," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 1, no. 2, pp. 37-44, 2019, doi: 10.47709/cnipc.v1i2.239.
- [3] C. L. Clayman, S. M. Srinivasan, R. S. Sangwan, C. L. Clayman, S. M. Srinivasan, and R. S. Sangwan, "ScienceDirect ScienceDirect K-means Clustering and Principal Components Analysis of Microarray Data L1000 Components Landmark Genes K-means Clustering and of Principal Analysis of Microarray Data of L1000 Landmark Genes," *Procedia Comput. Sci.*, vol. 168, no. 2019, pp. 97-104, 2020, doi: 10.1016/j.procs.2020.02.265.
- [4] Y. A. Wijaya, D. A. Kurniady, E. Setyanto, W. S. Tarihoran, D. Rusmana, and R. Rahim, "Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities," *TEM J.*, vol. 10, no. 3, pp. 1099-1103, 2021, doi: 10.18421/TEM103-13.
- [5] "https://docs.rapidminer.com/10.0/studio/operators/modeling/segmentation/k_means.html".
- [6] I. P. Mulyadi, "Jurnal Informatika Ekonomi Bisnis Klasterisasi Menggunakan Metode Algoritma K-Means dalam Meningkatkan Penjualan Tupperware," vol. 4, pp. 5-9, 2022, doi: 10.37034/infec.v4i4.164.
- [7] A. Nofiar, S. Defit, and Sumijan, "Penentuan Mutu Kelapa Sawit Menggunakan Metode K-Means Clustering," *J. KomtekInfo*, vol. 5, no. 3, pp. 1-9, 2019, doi: 10.35134/komtekinfoc.v5i3.26.
- [8] J. Santos-Pereira, L. Gruenwald, and J. Bernardino, "Top data mining tools for the healthcare industry," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 8, pp. 4968-4982, 2022, doi: 10.1016/j.jksuci.2021.06.002.
- [9] M. S. H. Ardani *et al.*, "A new approach to signal filtering method using K-means clustering and distance-based Kalman filtering," *Sens. Bio-Sensing Res.*, vol. 38, no. July, p. 100539, 2022, doi: 10.1016/j.sbsr.2022.100539.
- [10] I. Hadi, L. W. Santoso, and A. N. Tjondrowiguno, "Sistem Rekomendasi Film menggunakan User-based Collaborative Filtering dan K-modes Clustering," *J. Infra*, vol. 3, no. 1, pp. 18-21, 2020.
- [11] H. Syahputra, "Clustering Tingkat Penjualan Menu (Food and Beverage) Menggunakan Algoritma K-Means," *J. KomtekInfo*, vol. 9, pp. 29-33, 2022, doi: 10.35134/komtekinfoc.v9i1.274.