

Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Menganalisa Data Film Lokal (Indonesia) Populer

Risma Rito Pakpahan

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: rismarp20@gmail.com

Email Penulis Korespondensi: rismarp20@gmail.com

Submitted 20-02-2025; Accepted 28-02-2025; Published 28-02-2025

Abstrak

Film merupakan media audio visual yang terdiri dari gabungan potongan gambar yang menjadi satu kesatuan dan didalamnya terdapat pesan yang terkandung pada setiap alur ceritanya. Bioskop merupakan tempat pertunjukkan film yang dipertunjukkan dengan lampu yang disorot pada layar lebar. Pemilihan film yang dipertunjukkan merupakan faktor penting yang mempengaruhi tingkat jumlah penonton pada bioskop, akan tetapi pemilihan film yang sesuai dengan minat penonton tidaklah gampang, sehingga masalah ini berdampak pada tidak stabilnya jumlah penonton pada bioskop, bahkan dapat merugikan karena kurangnya penonton pada film yang di pertunjukkan. Solusi yang dapat dilakukan pada permasalahan ini adalah melakukan pengelompokan terhadap data film. Pengelompokan tersebut dilakukan untuk mengetahui film yang paling banyak penonton dan yang tidak banyak penontonnya. algoritma yang digunakan adalah Algoritma K-Means, dimana algoritma ini mampu melakukan pengelompokan data berdasarkan karakteristiknya dengan menentukan nilai klaster, kemudian data diolah pada setiap literasi hingga tidak ada nilai klaster yang berubah. Penerapan algoritma K-Means pada data film lokal (Indonesia) berhasil mengelompokkan data film dimana pada data film yang diolah dihasilkan data film lokal yang paling banyak penonton serta data film lokal yang kurang di minati oleh penonton. Sehingga pada penerapan metode K-Means memberikan informasi film yang dapat sering dipertunjukkan pada bioskop jumlah penonton bahkan menjadi lebih banyak.

Kata Kunci: Film; Bioskop; K-Means

Abstract

Film is an audiovisual medium composed of a series of images that form a unified whole, with messages embedded in its storyline. A cinema is a venue where films are shown using projected light on a large screen. The selection of films to be screened is a crucial factor influencing the number of moviegoers. However, choosing films that align with audience preferences is not easy, leading to unstable audience numbers and potential financial losses due to low viewership. A solution to this issue is to classify film data to identify which films attract the most and the least viewers. The algorithm used for this classification is the K-Means algorithm, which groups data based on its characteristics by determining cluster values and iteratively processing the data until the cluster values remain unchanged. The application of the K-Means algorithm to local (Indonesian) film data successfully categorizes films based on audience interest, identifying the most-watched local films and those with lower viewership. Thus, implementing the K-Means method provides insights into films that can be frequently screened in cinemas, potentially increasing audience numbers.

Keywords: Film; Cinema; K-Means

1. PENDAHULUAN

Hiburan merupakan salah satu hal yang dibutuhkan oleh manusia untuk mendapat suatu kesenangan tersendiri terhadap hiburan yang diminati, Salah satunya adalah film. Film adalah media audio visual yang terdiri dari potongan gambar yang disatukan menjadi kesatuan utuh dan memiliki kemampuan dalam menangkap realita sosial budaya serta memiliki pesan yang terkandung didalamnya dalam bentuk visual [1]. Film dalam kamus besar bahasa Indonesia, memiliki arti sebagai selaput tipis yang terbuat dari seluloid yang berfungsi sebagai tempat gambar negatif (yang dibuat potret) maupun gambar positif (yang dimainkan di bioskop). Selain itu, film juga diartikan sebagai lakon (cerita) gambar hidup [2]. Pada saat ini salah satu tempat populer untuk menonton film adalah bioskop. Bioskop pertama di Indonesia berdiri pada Desember 1900, di Jl Tanah Abang I, Jakarta Pusat, karcis kelas I harganya dua gulden (perak) dan harga karcis kelas dua setengah perak. Bioskop zaman dulu bermula di sekitar Lapangan Gambir (kini Monas) [1]. Bioskop merupakan Pertunjukan yang diperlihatkan dengan gambar (film) yang disorot menggunakan lampu sehingga dapat bergerak (berbicara) [3]. Salah satu tempat menonton film yaitu Cinema XXI Deli Park Medan.

Perkembangan dunia teknologi pada saat ini menjadi salah satu kendala yang terjadi di bioskop Cinema XXI Deli Park Medan dimana sebagian film yang dipertontonkan di bioskop dapat di akses atau ditonton oleh masyarakat melalui internet, salah satunya adalah Youtube. Permasalahan tersebut memberikan dampak pada bioskop Cinema XXI Deli Park Medan menjadi tidak stabil dimana kadang ramai penonton terkadang sepi. Hal ini terjadi dikarenakan tidak diketahui tingkat kepopuleran suatu film yang akan dipertunjukkan di bioskop.

Berdasarkan permasalahan diatas maka solusi yang dapat diambil adalah melakukan pemilihan pada persediaan film yang di pertontonkan kepada penonton, sehingga film yang ditampilkan hanya berupa film populer yang ramai penontonnya. Maka pada pemilihan tersebut dibutuhkan suatu metode yang dapat melakukan pemilihan pada film sehingga dapat menghasilkan suatu keputusan tentang film yang akan di pertontonkan kepada penonton. Pada penelitian ini metode yang diguakan adalah metode Clustering dengan Algoritma K-Means, Clustering merupakan metode yang digunakan untuk membagi rangkaian data menjadi beberapa kelompok berdasarkan kemiripan karakteristik yang telah ditentukan sebelumnya [4] sedangkan K-Means merupakan algoritma yang menetapkan nilai-nilai Cluster (k) secara random, untuk sementara nilai tersebut menjadi pusat dari Cluster yang biasa disebut centroid [5].

Penelitian yang dilakukan oleh Nayla Nur Afidaha, Masrukan (2023), mengatakan bahwa metode Cluster_ing dengan algoritma K-Means dapat digunakan untuk mengelompokkan data migrasi penduduk dengan cara mengklasifikasikan tiap kecamatan pada Cluster yang telah ditetapkan. Berdasarkan hasil Cluster_ing dengan algoritma K-Means antara perhitungan manual dengan aplikasi terdapat sedikit perbedaan dimana Cluster1 (C1), pada perhitungan manual diperoleh sebanyak 7 kecamatan meliputi Kecamatan Sumber, Bulu, Gunem, Sulang, Kaliori, Pancur, dan Sluke. Sedangkan pada aplikasi diperoleh sebanyak 10 kecamatan, meliputi Kecamatan Sumber, Bulu, Gunem, Sale, Sedan, Pamotan, Sulang, Kaliori, Pancur, dan Sluke. Cluster2 (C2), pada perhitungan manual diperoleh sebanyak 6 kecamatan, meliputi Kecamatan Sale, Sarang, Sedan, Pamotan, Kragan, dan Lasem. Sedangkan pada aplikasi diperoleh sebanyak 3 kecamatan, meliputi Kecamatan Sarang, Kragan, dan Lasem. Cluster3 (C3), baik perhitungan manual maupun aplikasi diperoleh hasil yang sama sebanyak 1 kecamatan yaitu Kecamatan Rembang [6].

Penelitian yang dilakukan oleh Yan Watequlis Syaifudin, RizkiAndi Irawan (2018), Metode K-Means memberikan hasil dimana pada implementasi Support Vector Machine (SVM) pada Analisis Sentimen bekerja dengan baik, rata-rata accuracy yang didapatkan oleh sistem sebesar 74,39%, dan Hasil K-Means Cluster_ing dipengaruhi dari nilai titik pusat Cluster (centroid) yang dan jumlah data yang digunakan. Selain itu perbedaan pengambilan data pusat Cluster awal yang digunakan juga akan mempengaruhi hasil akhir pengelompokkan [7].

Penelitian yang dilakukan oleh Zacky Yaser Malik Gumiwang, Ahmad Fahry Hamidy, Eva Yulia Puspaningrum (2023), mengatakan bahwa kondisi indikator kualitas film masih memiliki ketidakseimbangan karena ada beberapa kategori film yang memiliki kualitas buruk dan kategori film yang tidak jelas keterangannya. Dengan 5 Cluster menggunakan metode Cluster_ing K-means, diperoleh Cluster 0 dengan kualitas film yang buruk, Cluster 1 dengan kualitas film sangat buruk, Cluster 2 dengan kualitas film baik, Cluster 3 dengan kualitas film sangat baik sedangkan Cluster 4 tidak jelas keterangannya [8].

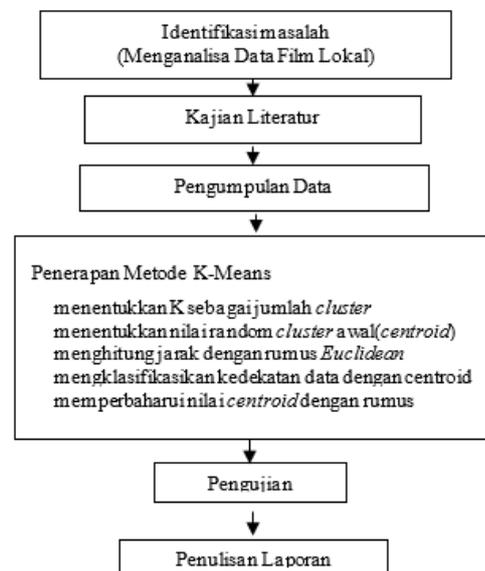
Penelitian yang dilakukan oleh Syarifah Iin Safitri, Cucu Suhery, Syamsul Bahri (2021), mengatakan bahwa pengujian yang telah dilakukan pada sistem menghasilkan presentase keberhasilan Algoritma K-Means pada maskapai Garuda Indonesia sebesar 71,8%, Air Asia sebesar 52,8% dan Lion Air sebesar 62,1% dan Faktor yang dapat mempengaruhi akurasi pada algoritma K-Means yaitu penentuan pusat cluster atau inisialisasi centroid awal yang dilakukan secara acak [9].

Berdasarkan permasalahan yang telah dijelaskan, pemilihan film yang sesuai dengan minat penonton menjadi faktor penting dalam meningkatkan jumlah penonton di bioskop. Namun, menentukan film yang tepat tidaklah mudah, sehingga berdampak pada ketidakstabilan jumlah penonton. Oleh karena itu, diperlukan metode analisis data yang efektif untuk mengelompokkan film berdasarkan jumlah penonton. Penulis tertarik melakukan penelitian mengenai penerapan metode clustering dengan algoritma K-Means dalam menganalisis data film lokal (Indonesia) yang populer di Cinema XXI Deli Park Medan. Dengan metode ini, diharapkan dapat diperoleh informasi yang lebih akurat mengenai film yang diminati dan kurang diminati oleh penonton.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam pelaksanaan penelitian, disiapkan sebuah kerangka kerja penelitian yang berisi langkah-langkah pelaksanaan kegiatan untuk mempermudah proses penelitian yang berlangsung agar lebih teratur dan sistematis. Adapun tahapan kerja penelitian yang dilakukan dapat dilihat pada gambar dibawah ini :



Gambar 1. Kerangka Penelitian

Adapun penjelasan dari kerangka kerja penelitian pada gambar 1 adalah sebagai berikut :

1. Identifikasi Masalah
Identifikasi masalah adalah tahapan awal untuk menjelaskan sebuah permasalahan dalam menganalisa data film lokal yang akan di lakukan klasifikasi.
2. Kajian Literatur
Kajian literatur merupakan riset dengan membaca berbagai jurnal, artikel ilmiah, buku, atau *website* mengenai permasalahan yang dihadapi serta topik yang ingin diteliti. Dalam mengkaji sebuah literatur harus melalui sumber yang jelas dan dapat dipertanggungjawabkan.
3. Pengumpulan Data
Pengumpulan data penelitian bisa berupa sampel data yang kemudian akan diolah sesuai metode yang diterapkan.
4. Implementasi Metode K-Means
Implementasi metode K-Means merupakan proses klasifikasi dari data film yang telah dilakukan analisa sebelumnya sehingga memperoleh hasil.
5. Pengujian Sistem
Pengujian sistem (*testing*) dilakukan untuk menguji hasil klasifikasi menggunakan metode K-Means ke dalam sistem untuk dilakukan pengujian.
6. Penulisan Laporan
Penulisan laporan dilakukan untuk mendokumentasi seluruh kegiatan penelitian dalam bentuk skripsi yang nantinya juga dibuat dalam bentuk artikel ilmiah yang akan dipublikasikan.

2.2 Film

Film adalah karya cipta seni dan budaya yang merupakan media komunikasi massa pandang-dengar yang dibuat berdasarkan asas sinematografi dengan direkam pada pita seluloid, pita video, piringan video dan atau bahan hasil penemuan teknologi lainnya dalam segala bentuk, jenis dan ukuran melalui proses kimiawi, proses elektronik atau proses lainnya, dengan atau tanpa suara, yang dapat dipertunjukkan dan atau ditayangkan dengan sistem proyeksi mekanik, eletronik, dan atau lainnya [10]. Film dalam kamus besar bahasa indonesia, memiliki arti sebagai selaput tipis yang terbuat dari seluloid yang berfungsi sebagai tempat gambar negatif (yang dibuat potret) maupun gambar positif (yang dimainkan di bioskop). Selain itu, film juga diartikan sebagai lakon (cerita) gambar hidup [2]. Berdasarkan pengertiannya diatas dapat disimpulkan bahwa film merupakan sebuah hiburan visual yang terdiri dari potongan gambar, suara, yang disatukan dalam satu kesatuan dimana didalamnya mengandung seni dan budaya serta memiliki pesan yang disampaikan pada setiap film.

2.3 Bioskop

Kata bioskop berasal dari kata bioscoop (bahasa Belanda yang juga berasal dari Bahasa Yunani) adalah tempat bagi masyarakat untuk menonton pertunjukan film dengan menggunakan layar lebar. Gambar film diproyeksikan ke layar menggunakan proyektor. Bioskop merupakan Pertunjukan yang diperlihatkan dengan gambar (film) yang disorot menggunakan lampu sehingga dapat bergerak (berbicara) [3]. Bioskop adalah pertunjukan yang diperlihatkan dengan gambar (film) yang disorot sehingga bergerak (berbicara) [11]. Secara khusus “bioskop” dapat diartikan sebagai tempat bercengkrama (rendevous) bagi pembuat (sineas) dengan penggemar/pecinta seni film dan atau film seni [12]. Berdasarkan pengertiannya diatas dapat disimpulkan bahwa bioskop merupakan tempat untuk menonton film yang diperlihatkan dalam bentuk layar lebar menggunakan proyektor.

2.4 Algoritma K-Means

K-Means merupakan salah satu algoritma pengelompokan yang melakukan partisi set data kedalam sejumlah cluster yang sudah ditetapkan di awal dengan mengambil nilai minimal dari dari jarak titik data terhadap cluster. Proses perhitungan jarak dihitung menggunakan rumus euclidian distance [13]. Diberikan satu set pengamatan (x_1, x_2, \dots, x_n) , di mana setiap observasi adalah d-dimensi vector nyata, pengelompokan K-Means bertujuan untuk mempartisin pengamatan kedalam k himpunan $(k \leq n)$ $S = \{S_1, S_2, \dots, S_k\}$ sehingga menjadi untuk meminimalkan Jumlah Kuadrat Dalam-Cluster [114].

Berdasarkan pengertian diatas dapat disimpulkan bahwa algoritma K-Means melakukan pengelompokkan kedalam sejumlah cluster. Berikut merupakan cara kerja algoritma K-Means [13]:

1. Menentukan K sebagai jumlah *cluster*.
2. Membangkitkan nilai random untuk pusat *cluster* awal (*centroid*) sebanyak k.
3. Menghitung jarak setiap data input terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean* (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidian Distance* :

$$d(x_i, u_j) = \sqrt{\sum (x_i - u_j)^2} \quad (1)$$

keterangan :

x_i : data kriteria

u_j : *centroid* pada *cluster* ke-j

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).

- Memperbaharui nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$uj(t + 1) = \frac{1}{nsj} \sum_{j \in sj} x_j \quad (2)$$

Keterangan :

$uj(t + 1)$: *centroid* baru pada literasi ke (t+1)

nsj : banyak data pada *cluster* sj.

- Melakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah. Jika langkah 6 telah terpenuhi, maka nilai pusat *cluster* (uj) pada literasi terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Data

Data film lokal (Indonesia) merupakan sekumpulan film lokal yang ditayangkan di bioskop Cinema XXI Deli Park Medan, dimana data film lokal tersebut terusun secara acak sehingga tidak diketahui film lokal (Indonesia) yang paling banyak ditonton diantara film lokal (Indonesia) lainnya. Berikut merupakan sampel data yang digunakan pada penelitian ini, dapat dilihat pada tabel 4.1 dibawah :

Tabel 1. Sampel Data

NO	Judul Film	Jumlah Penonton	Jumlah Tayang	GENRE
1	Buya Hamka	1250	19	Drama
2	Hello Ghost	613	11	Horor/Komedi
3	Bayi Ajaib	434	19	Horor
4	Waktu Maghrib	1503	9	Horor
5	perjanjian gaib	352	9	Horor/Komedi
6	Hidayah	333	12	Horor
7	Kutukan Sembilan Setan	244	8	Horor
8	Khanzab	1132	19	Horor
9	Jalan yang Jauh, Jangan Lupa Pulang	861	2	Drama
10	Hati Suhita	502	25	Drama
11	Bismillah Kunikahi Suamimu	273	23	Romantis
12	Sewu Dino	3360	19	Horor
13	Spirit Doll	279	1	Horor
14	Mangkujiwo 2	555	26	Horor
15	Sosok Ketiga	1042	22	Horor

Sampel data diatas kemudian dilakukan perubahan untuk setiap label, dengan tujuan untuk mempersingkat label dari setiap data. Berikut merupakan atribut untuk setiap data.

Tabel 2. Sampel Data

DATA	X	Y
F1	1250	19
F2	613	11
F3	434	19
F4	1503	9
F5	352	9
F6	333	12
F7	244	8
F8	1132	19
F9	861	2
F10	502	25
F11	273	23
F12	3360	19
F13	279	1
F14	555	26
F15	1042	22

Keterangan:

F1 = merupakan judul film (F) dan angka 1 merupakan data judul film ke-1

X = jumlah penonton

Y = jumlah tayang

3.2 Penerapan Algoritma K-Means

1. Menentukan jumlah kluster
2. Menentukan titik pusat kluster awal (secara acak)

Tabel 3. Menentukan Titik Pusat Kluster Awal

DATA	X	Y	C1	C2
F1	1250	19		
F2	613	11		
F3	434	19		
F4	1503	9		
F5	352	9		
F6	333	12		
F7	244	8		
F8	1132	19		
F9	861	2		
F10	502	25		
F11	273	23		
F12	3360	19		
F13	279	1		
F14	555	26		
F15	1042	22		

Tabel 4. Titik Pusat Kluster Awal

Data	Kluster	X	Y
F1	C1	1250	19
F7	C2	244	8

Keterangan Kluster :

C1 = Populer

C2 = Tidak Populer

3. Menghitung jarak data dengan kluster dengan rumus :

$$d(x_i, u_j) : \sqrt{\sum (x_i - u_j)^2}$$

Literasi 1:

Menghitung jarak data dengan C1

$$d(F1, C1) = \sqrt{(1250 - 0)^2 + (19 - 19)^2}$$

$$d(F1, C1) = 0$$

$$d(F2, C1) = \sqrt{(613 - 0)^2 + (11 - 19)^2}$$

$$d(F2, C1) = 637.0502$$

$$d(F3, C1) = \sqrt{(434 - 0)^2 + (19 - 19)^2}$$

$$d(F3, C1) = 816$$

$$d(F4, C1) = \sqrt{(1503 - 0)^2 + (9 - 19)^2}$$

$$d(F4, C1) = 253.1976$$

$$d(F5, C1) = \sqrt{(352 - 0)^2 + (9 - 19)^2}$$

$$d(F5, C1) = 898.0557$$

$$d(F6, C1) = \sqrt{(333 - 0)^2 + (12 - 19)^2}$$

$$d(F6, C1) = 917.0267$$

$$d(F7, C1) = \sqrt{(244 - 0)^2 + (8 - 19)^2}$$

$$d(F7, C1) = 1006.06$$

$$d(F8, C1) = \sqrt{(1132 - 0)^2 + (19 - 19)^2}$$

$$d(F8, C1) = 118$$

$$d(F9, C1) = \sqrt{(861 - 0)^2 + (2 - 19)^2}$$

$$d(F9, C1) = 389.3713$$

$$d(F10, C1) = \sqrt{(502 - 0)^2 + (25 - 19)^2}$$

$$d(F10, C1) = 748.0241$$

$$d(F11, C1) = \sqrt{(273 - 0)^2 + (23 - 19)^2}$$

$$d(F11, C1) = 977.0082$$

$$d(F12, C1) = \sqrt{(3360 - 0)^2 + (19 - 19)^2}$$

$$d(F12, C1) = 2110$$

$$d(F13, C1) = \sqrt{(279 - 0)^2 + (1 - 19)^2}$$

$$d(F13, C1) = 971.1668$$

$$d(F14, C1) = \sqrt{(555 - 0)^2 + (26 - 19)^2}$$

$$d(F14, C1) = 695.0353$$

$$d(F15, C1) = \sqrt{(1042 - 0)^2 + (22 - 19)^2}$$

$$d(F15, C1) = 208.0216$$

Menghitung jarak data dengan C2

$$d(F1, C2) = \sqrt{(1250 - 244)^2 + (19 - 8)^2}$$

$$d(F1, C2) = 1006.06$$

$$d(F2, C2) = \sqrt{(613 - 244)^2 + (11 - 8)^2}$$

$$d(F2, C2) = 369.0122$$

$$d(F3, C2) = \sqrt{(434 - 244)^2 + (19 - 8)^2}$$

$$d(F3, C2) = 190.3182$$

$$d(F4, C2) = \sqrt{(1503 - 244)^2 + (9 - 8)^2}$$

$$d(F4, C2) = 1259$$

$$d(F5, C2) = \sqrt{(352 - 244)^2 + (9 - 8)^2}$$

$$d(F5, C2) = 108.0046$$

$$d(F6, C2) = \sqrt{(333 - 244)^2 + (12 - 8)^2}$$

$$d(F6, C2) = 89.08984$$

$$d(F7, C2) = \sqrt{(244 - 244)^2 + (8 - 8)^2}$$

$$d(F7, C2) = 0$$

$$d(F8, C2) = \sqrt{(1132 - 244)^2 + (19 - 8)^2}$$

$$d(F8, C2) = 888.0681$$

$$d(F9, C2) = \sqrt{(861 - 244)^2 + (2 - 8)^2}$$

$$d(F9, C2) = 617.0292$$

$$d(F10, C2) = \sqrt{(502 - 244)^2 + (25 - 8)^2}$$

$$d(F10, C2) = 258.5595$$

$$d(F11, C2) = \sqrt{(273 - 244)^2 + (23 - 8)^2}$$

$$d(F11, C2) = 32.64966$$

$$d(F12, C2) = \sqrt{(3360 - 244)^2 + (19 - 8)^2}$$

$$d(F12, C2) = 3116.019$$

$$d(F13, C2) = \sqrt{(279 - 244)^2 + (1 - 8)^2}$$

$$d(F13, C2) = 35.69314$$

$$d(F14, C2) = \sqrt{(555 - 244)^2 + (26 - 8)^2}$$

$$d(F14, C2) = 311.5205$$

$$d(F15, C2) = \sqrt{(1042 - 244)^2 + (22 - 8)^2}$$

$$d(F15, C2) = 798.1228$$

Berikut merupakan hasil perhitungan jarak pada literasi 1, kemudian menentukan nilai minimum untuk mencari kluster

Tabel 5. Literasi 1

DATA	X	Y	C1	C2	MIN
F1	1250	19	0	1006.06	0
F2	613	11	637.0502	369.0122	369.0122
F3	434	19	816	190.3182	190.3182
F4	1503	9	253.1976	1259	253.1976
F5	352	9	898.0557	108.0046	108.0046
F6	333	12	917.0267	89.08984	89.08984
F7	244	8	1006.06	0	0
F8	1132	19	118	888.0681	118
F9	861	2	389.3713	617.0292	389.3713
F10	502	25	748.0241	258.5595	258.5595
F11	273	23	977.0082	32.64966	32.64966
F12	3360	19	2110	3116.019	2110
F13	279	1	971.1668	35.69314	35.69314
F14	555	26	695.0353	311.5205	311.5205
F15	1042	22	208.0216	798.1228	208.0216

4. Mengelompokkan data berdasarkan jarak terkecil

Tabel 6. Mengelompokkan Jarak Terkecil

DATA	X	Y	C1	C2	MIN	KLASTER
F1	1250	19	0	1006.06	0	1
F2	613	11	637.0502	369.0122	369.0122	2
F3	434	19	816	190.3182	190.3182	2
F4	1503	9	253.1976	1259	253.1976	1
F5	352	9	898.0557	108.0046	108.0046	2
F6	333	12	917.0267	89.08984	89.08984	2
F7	244	8	1006.06	0	0	2
F8	1132	19	118	888.0681	118	1
F9	861	2	389.3713	617.0292	389.3713	1
F10	502	25	748.0241	258.5595	258.5595	2
F11	273	23	977.0082	32.64966	32.64966	2
F12	3360	19	2110	3116.019	2110	1
F13	279	1	971.1668	35.69314	35.69314	2
F14	555	26	695.0353	311.5205	311.5205	2
F15	1042	22	208.0216	798.1228	208.0216	1

5. Memperbaharui titik pusat kluster

Rumus :

$$uj(t + 1) = \frac{1}{ns_j} \sum_{j \in s_j} x_j$$

Menghitung titik pusat kluster C1

$$uj(c1x) = \frac{1250+1503+1132+861+3360+1042}{6}$$

$$uj(c1x) = 1524.667$$

$$uj(c1y) = \frac{19+9+19+2+19+22}{6}$$

$$uj(c1y) = 15$$

Menghitung titik pusat kluster C2

$$uj(c2x) = \frac{613+434+352+333+244+502+273+555}{9}$$

$$uj(c2x) = 398.3333$$

$$uj(c2y) = \frac{11+19+9+12+8+25+23+1+26}{9}$$

$$uj(c2y) = 14.88889$$

Berikut merupakan titik pusat klaster yang baru.

Tabel 7. Titik Pusat Klaster

C1	1524.667	15
C2	398.3333	14.88889

- Melakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap klaster tidak ada yang berubah. Jika langkah 6 telah terpenuhi, maka nilai pusat klaster pada literasi terakhir akan digunakan sebagai parameter untuk menentukan klasifikasi data. Maka lakukan hingga literasi ke 7.
- Berdasarkan hasil pada setiap literasi yang telah dilakukan perubahan pada setiap literasi dapat dilihat pada tabel berikut :

Tabel 8. Hasil dari Setiap Literasi

NO	Judul Film	Jumlah Penonton	Jumlah Tayang	Literasi 1	Literasi 2	Literasi 3	Literasi 4	Literasi 5	Literasi 6	Literasi 7
1	Buya Hamka	1250	19	1	1	1	1	2	2	2
2	Hello Ghost	613	11	2	2	2	2	2	2	2
3	Bayi Ajaib	434	19	2	2	2	2	2	2	2
4	Waktu Maghrib	1503	9	1	1	1	1	1	2	2
5	perjanjian gaib	352	9	2	2	2	2	2	2	2
6	Hidayah	333	12	2	2	2	2	2	2	2
7	Kutukan Sembilan Setan	244	8	2	2	2	2	2	2	2
8	Khanzab	1132	19	1	1	1	2	2	2	2
9	Jalan yang Jauh, Jangan Lupa Pulang	861	2	1	2	2	2	2	2	2
10	Hati Suhita	502	25	2	2	2	2	2	2	2
11	Bismillah Kunikahi Suamimu	273	23	2	2	2	2	2	2	2
12	Sewu Dino	3360	19	1	1	1	1	1	1	1
13	Spirit Doll	279	1	2	2	2	2	2	2	2
14	Mangkujiwo 2	555	26	2	2	2	2	2	2	2
15	Sosok Ketiga	1042	22	1	1	2	2	2	2	2

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa analisis yang dilakukan berhasil membantu dalam mengolah data film lokal (Indonesia) yang dipertunjukkan di bioskop Cinema XXI Deli Park Medan. Penerapan metode clustering dengan algoritma K-Means juga terbukti efektif dalam membantu bioskop menentukan film lokal yang akan ditayangkan. Hasil pengujian menunjukkan bahwa proses clustering berlangsung selama tujuh literasi, di mana iterasi berhenti pada literasi keenam karena pada literasi ketujuh tidak terjadi perubahan nilai klaster, dengan klaster pertama berjumlah satu dan klaster kedua berjumlah empat belas. Film merupakan media audiovisual yang terdiri dari gabungan potongan gambar yang menjadi satu kesatuan dan memiliki pesan dalam setiap alur ceritanya. Pemilihan film yang akan ditayangkan menjadi faktor penting dalam menentukan jumlah penonton di bioskop, namun memilih film yang sesuai dengan minat penonton bukanlah hal yang mudah. Hal ini dapat menyebabkan ketidakstabilan jumlah penonton dan bahkan dapat merugikan pihak bioskop jika film yang ditayangkan tidak diminati. Oleh karena itu, solusi yang dapat diterapkan adalah melakukan pengelompokan data film menggunakan algoritma K-Means. Algoritma ini mampu mengelompokkan film berdasarkan karakteristik tertentu dengan menentukan nilai klaster, lalu mengolah data hingga tidak terjadi perubahan klaster. Penerapan K-Means berhasil mengelompokkan film lokal yang paling diminati dan yang kurang diminati, sehingga dapat menjadi acuan dalam menentukan film yang akan sering diputar untuk meningkatkan jumlah penonton.

REFERENCES

- [1] M. S. Muhammad Ali Mursid Alfathoni, M.Sn., Dani Manesah, *Pengantar Teori Film*. Deepublish, 2020.
- [2] *kamus besar bahasa indonesia*. balai pustaka, 1990.
- [3] Andi, *Pembelajaran Bahasa Indonesia Yang Efektif di Sekolah Dasar*. Jakarta: Departemen Pendidikan Nasional, 2019.
- [4] Mardalius, "Pemanfaatan Rapid Miner Studio8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means.," *J. Teknol. dan Sist. Inf.*, vol. 4(2), pp. 123–132, 2018.

- [5] R. T. Vlandari, *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta, 2017.
- [6] N. Nur Afidah, "Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokkan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," *Prism. Pros. Semin. Nas. Mat.*, vol. 6, pp. 729–738, 2023, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [7] Y. W. Syaifudin and R. A. Irawan, "Implementasi Analisis Clustering Dan Sentimen Data Twitter Pada Opini Wisata Pantai Menggunakan Metode K-Means," *J. Inform. Polinema*, vol. 4, no. 3, p. 189, 2018, doi: 10.33795/jip.v4i3.205.
- [8] M. A. K.- Means, Z. Yaser, M. Gumiwang, A. F. Hamidy, and E. Y. Puspaningrum, "Analisis Data Mining Untuk Clustering Data Film Dengan," *J. Manaj. Inform. Jayakarta*, vol. 7, pp. 52–56, 2023.
- [9] S. B. Syarifah Iin Safitri, Cucu Suhery, "IMPLEMENTASI ALGORITMA K – MEANS UNTUK CLUSTERING SENTIMEN PADA OPINI KUALITAS PELAYANAN JASA PENERBANGAN," *J. Komput. dan Apl.*, vol. 09, pp. 186–197, 2021.
- [10] NKRI, "UNDANG-UNDANG REPUBLIK INDONESIA," in *PERFILMAN*, 1992, p. 8.
- [11] KBBI, "Kamus Besar Bahasa Indonesia," *Online*, 2016.
- [12] (Tjasmadi, "Sejarah Bioskop," *Sinemat. Indones.*, vol. 1, p. 1, 1992.
- [13] E. Yulian, "Text Mining dengan KMeans Clustering pada Tema LGBT dalam Arsip Tweet Masyarakat Kota Bandung," *MANTIK*, vol. 4, p. 1, 2018.
- [14] N. R. Himmah, "Implementasi Algoritma K-Means Untuk Pengelompokan Siswa Berdasarkan Nilai Akademik (Studi Kasus Mtsn Gresik)," *Undergrad. Thesis*, 2019.