

SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *K-MEANS* *CLUSTERING*

(MOVIE RECOMMENDATION SYSTEM USING *K-MEANS CLUSTERING* *METHOD*)

Herdita Mutiasari¹, Tito Waluyo Purboyo², Ratna Astuti Nugrahaeni³

^{1,2,3} Universitas Telkom, Bandung

herditaamutiasari@student.telkomuniversity.ac.id¹, titowaluyo@telkomuniversity.ac.id²,
ratnaan@telkomuniversity.ac.id³

Abstrak

Ketika kita berbicara mengenai tayangan hiburan khususnya film, ada kategori yang dibuat untuk membuat segmentasi penonton seperti batasan usia dan pemilihan genre dalam setiap film yang di produksi. Sehingga dibutuhkan sistem rekomendasi untuk memilih genre film yang sesuai selera dan usia penonton. Sistem rekomendasi film adalah sistem yang dibangun agar pemilihan film lebih mudah untuk para penggemar film. Sistem rekomendasi ini bertujuan agar para penggemar film dapat memilih film melalui rekomendasi rating atau pemilihan genre dan batasan usia yang telah dipilih oleh para pengguna. Sistem rekomendasi ini hanya berbasis website dan menggunakan metode K-Means Clustering.

Pada Tugas Akhir ini menggunakan metode K-Means Clustering dengan jumlah cluster sebanyak 2. Kelompok nilai berdasarkan rating merupakan Cluster 1 dan kelompok nilai berdasarkan genre merupakan Cluster 2. Implementasi tugas akhir ini menggunakan sebuah aplikasi berbasis website dengan Bahasa pemrograman PHP. Pengujian website menggunakan metode alpha beta sehingga meminimalisir adanya kesalahan di sisi implementasi. Hasil pengujian dilakukan dengan 3,4,5 preferences genre, dari hasil pengujian 4 preferences genre menghasilkan nilai error yang lebih kecil yaitu 0,44%.

Kata kunci : *Sistem Rekomendasi, Film, K-Means Clustering, Website.*

Abstract

When we talk about entertainment shows, especially films, there are categories made to segment the audience such as age restrictions and genre selection in each film produced. So that a recommendation system is needed to choose a film genre that suits the taste and age of the audience. The movie recommendation system is a system built to make movie selection easier for movie fans. This recommendation system is intended so that film fans can choose films through rating recommendations or selection of genres and age restrictions that have been selected by users. This recommendation system is only website-based and uses the K-Means Clustering method.

In this final project using the K-Means Clustering method with the number of clusters as much as 2. The value group based on the rating is Cluster 1 and the value group based on the genre is Cluster 2. The implementation of this final project uses a website-based application with the PHP programming language. Website testing uses the alpha beta method so as to minimize errors on the implementation side. The results of the test are carried out with 3,4,5 genre preferences, from the results of the 4 genre preferences test, it produces a smaller error value of 0.44%.

Key Word : *Sistem Rekomendasi, Film, K-Means Clustering, Website.*

1. Pendahuluan

Pada era sekarang dunia teknologi dan perindustrian sudah sangat pesat kemajuannya. Salah satu contoh perindustrian yang sangat pesat kemajuannya yaitu perfilman. Semakin banyaknya hasil produksi film baik

dari dalam negeri maupun luar negeri. Terlalu banyak genre dari film yang ada di era sekarang, dari mulai film action hingga animasi yang menggunakan CG (Computer Graphics). Film juga berpengaruh pada siapa yang menontonnya dan film apa yang di tontonnya. Seperti menonton film dari luar negeri, user dapat sambil belajar menggunakan Bahasa Inggris dan Bahasa lainnya[1].

Pada dunia perfilman ada kategori yang dibuat untuk membuat segmentasi penonton seperti batasan usia dan pemilihan genre dalam setiap film yang di produksi. Sehingga dibutuhkan sistem rekomendasi untuk memilih genre film yang sesuai selera dan usia penonton.

Hingga saat ini penelitian mengenai sistem rekomendasi film sudah banyak dilakukan mengingat semakin meningkatnya penikmat film, baik dari luar negeri maupun dalam negeri, baik anak kecil hingga dewasa. Tetapi masih belum terlalu spesifik dari tiap film, dalam penelitian sebelumnya pun user yang digunakan masih terlalu sedikit[2]. Dan belum ada perekomendasi film menggunakan rating.

Dengan adanya permasalahan tersebut, pada tugas akhir ini dibuat sistem rekomendasi menggunakan metode K-Means Clustering berbasis website. Sistem rekomendasi akan memunculkan genre film yang dipilih oleh user atau rekomendasi dari genre film yang sering dilihat. User juga dapat memilih film dari nilai rating yang sudah ada. Clustering merupakan teknik pengelompokan data berdasarkan kemiripan dalam satu cluster. Pengelompokan sangat cocok untuk sistem rekomendasi ini, agar lebih mudah menentukan film yang ingin ditonton.

2. Dasar Teori

2.1 Clustering

Clustering adalah pengelompokan atau klasifikasi suatu data yang sama untuk menjadi suatu kelompok atau cluster yang berbeda untuk menjadikan suatu kelompok yang baru. Clustering sangat berguna untuk menemukan group atau kelompok yang tidak dikenal dalam suatu data dan untuk mengurangi jumlah data yang besar dengan memberikan kategori atau kelompok data yang memiliki kesamaan yang tinggi[7].

Metode clustering secara umum dibagi menjadi dua, yaitu hierarchical clustering dan non-hierarchical clustering. Hierarchical clustering merupakan suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Proses itu dilakukan berulang kali hingga cluster akan membentuk semacam pohon yang memiliki hierarchi (tingkatan) yang jelas antar objek. Sedangkan non-hierarchical clustering merupakan suatu metode pengelompokan data yang dimulai dengan menentukan jumlah cluster terlebih dahulu, setelah itu akan dilakukan proses clustering tanpa mengikuti hierarki[8].

2.2 K-Means Clustering

K-Means Clustering merupakan metode non-hierarchical clustering yang mengelompokkan data dalam bentuk cluster atau kelompok. Data akan dikelompokkan didalam cluster yang memiliki karakteristik yang sama, sehingga data yang sudah dikelompokkan memiliki variasi bentuk yang kecil.

Algoritma K-Means Clustering memiliki kelebihan, yaitu algoritma ini dinilai cukup efisien, yang ditunjukkan dengan kompleksitasnya. Serta algoritma ini akan berhenti dalam kondisi optimum local[8].

Selain kelebihan, algoritma K-Means Clustering memiliki kekurangan yaitu hasil pengelompokan akan bergantung pada acuan jumlah K cluster yang dan inialisasi nilai centroid awal yang digunakan.

Terdapat langkah-langkah dalam penggunaan algoritma K-Means Clustering, yaitu sebagai berikut : [9]

1. Menentukan jumlah cluster (K) yang akan dibentuk.
2. Inialisasi nilai centroid awal untuk masing masing cluster.
3. Hitung jarak antar objek ke masing masing centroid dengan menggunakan metode Euclidian Distance yang ada dibawah ini. Setelah dilakukan perhitungan setiap objek dimasukan ke dalam centroid dengan nilai centroid terdekat/minimum.

$$d_{ik} = \sqrt{\sum_{j=1}^m (X_{ij} - C_{kj})^2}$$

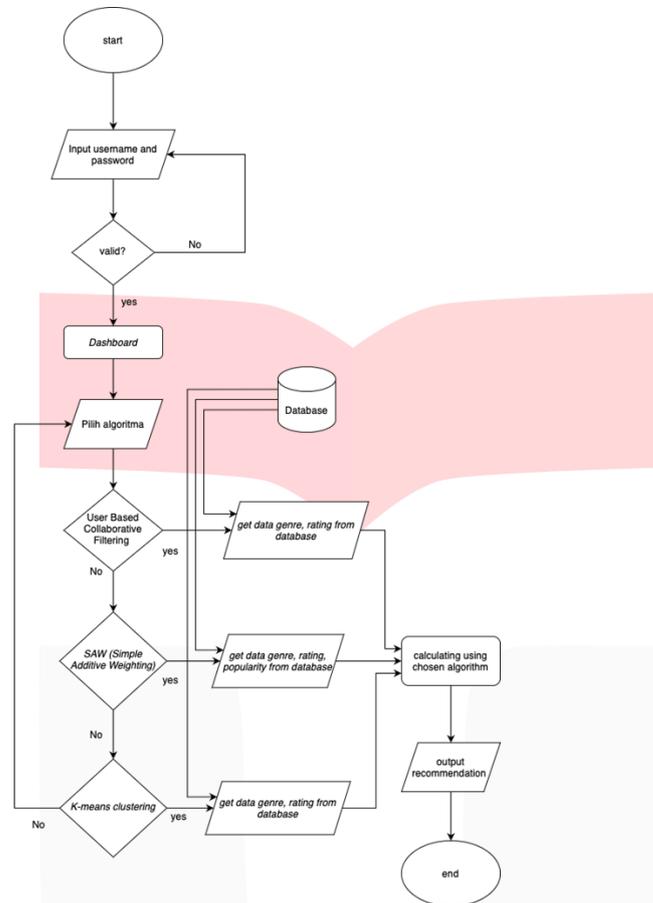
Diketahui:

- Jarak antara data ke-i dengan centroid tiap cluster
 - m adalah jumlah atribut
 - xij adalah data ke-i
 - ckj adalah data pusat tiap cluster.
4. Setiap objek pada cluster tersebut akan dihitung menjadi nilai centroid baru.
 5. Ulangi langkah 3 hingga posisi centroid sudah tidak mengalami perubahan.

3. Perancangan

3.1. Perancangan Sistem

Pada sistem ini akan dibangun sebuah aplikasi website untuk merekomendasikan film. Gambar dibawah menunjukkan alur proses sistem yang dibangun.



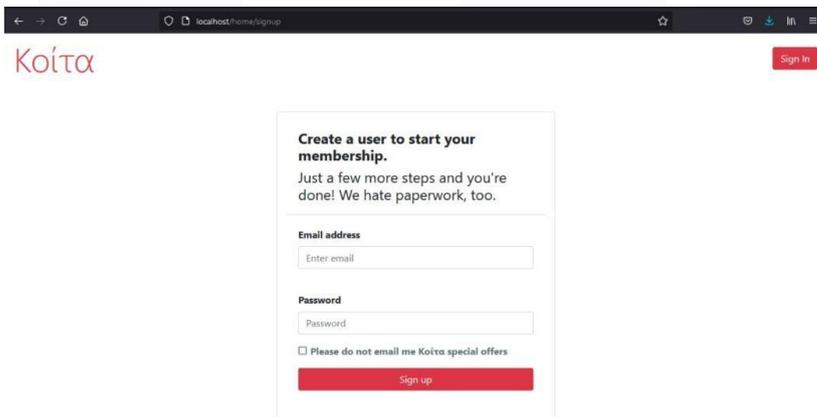
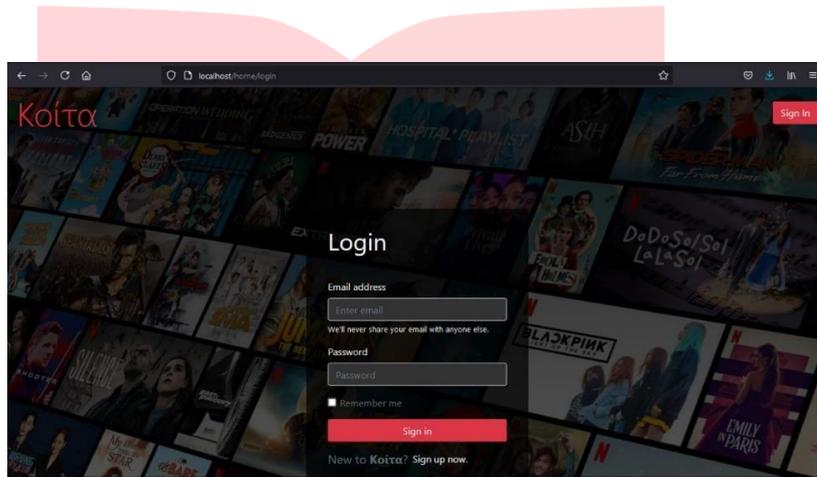
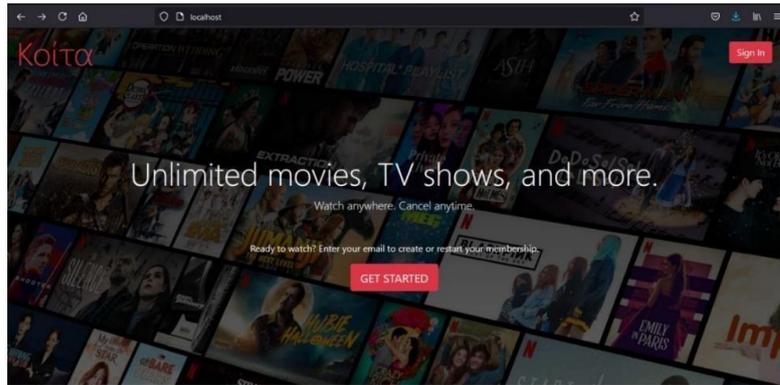
Proses berjalan diawali dengan user mengakses homepage. Selanjutnya user melakukan login untuk masuk kedalam dashboard. User yang dapat login yang sudah memiliki akun sebelumnya. Jika user belum memiliki akun maka harus melakukan registrasi terlebih dahulu. Saat user sudah masuk kedalam dashboard maka user akan diperintah untuk memilih salah satu algoritma yang telah tersedia. Setelah memilih algoritma, user menuju halaman algoritma. Setelah masuk ke halaman algoritma, untuk user baru akan diperintah untuk memilih sistem rekomendasi yang diinginkan seperti rating atau genre.

Secara umum pada aplikasi ini memiliki fitur lima menu utama yang dapat digunakan oleh user, yaitu :

- Menu login user yang sudah memiliki akun dapat langsung mengisikan username dan password untuk melakukan proses login dan masuk ke menu dashboard.
- Menu Register atau sign up dibuat untuk user membuat akun.
- Menu dashboard-home dibuat untuk user melihat film dari yang terbaru hingga yang paling populer.
- Menu dashboard-recommendation akan meminta user untuk memilih algoritma yang akan digunakan, menampilkan informasi film-film yang sesuai dengan kriteria yang sesuai dengan algoritma yang dipilih.
- Menu Algoritma K-Means Genre menampilkan rekomendasi film berdasarkan genre dari yang paling besar nilai bobotnya.
- Menu Algoritma K-Means Rating menampilkan rekomendasi film berdasarkan rating dari yang paling besar nilai bobotnya.

3.2. Antarmuka yang Dibangun

Berikut merupakan implementasi antarmuka yang dibangun.



Koita

Just a few more steps and you're done! We hate paperwork, too.

Please select 3 Preferences Genre that you like

Choose genre 1

Choose genre 2

Choose genre 3

Save

Hi!, Start your journey.

Watch your favorite movie here. All about yours, watch on your Smart TVs, Playstation, Xbox, Chromecast, Apple TV, Blu-ray players, and many more.

RECENTLY ADDED

Zulawski on Zulawski
A portrait of film director
ZULAWSKI ON ZULAWSKI
shows the different layers of
the personality of the artist
and does not shy away from
contrasts.
Release Date: 01-Jan-2000

Le Petit Poucet
PANTIN
la
dirouette
Release Date: 01-Jan-2000

A Constant Forge
One of the great mavericks
of cinema, John Cassavetes
has earned the reputation as
the father of American
independent movies. The
actor-turned-filmmaker
invented a realist...
Release Date: 01-Jan-2000

Ian Fleming: 007's
Creator
John Cork and his
colleagues from the Ian
Fleming Estate, who
have created the award
winning documentaries for most
of MCM's 007 DVDs, provide a
fine overview...
Release Date: 01-Jan-2000

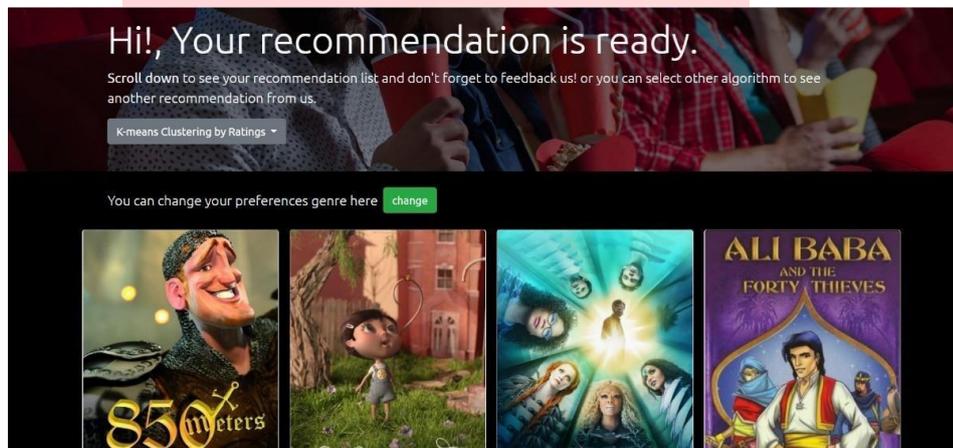
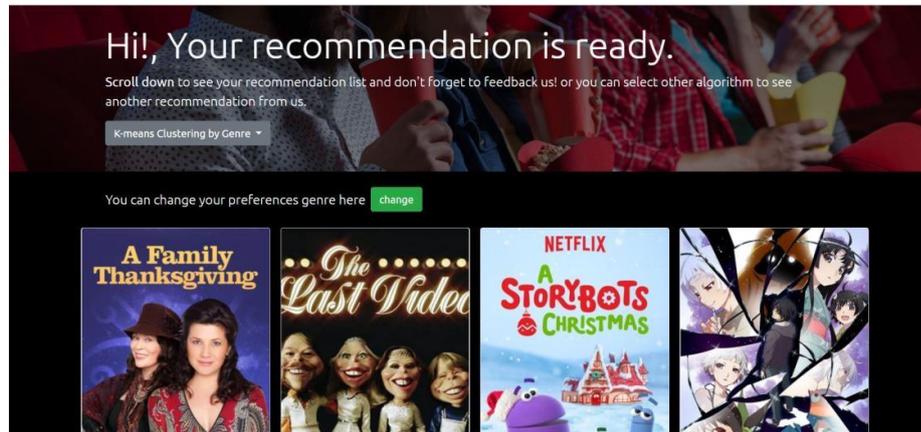
Koita Home Recommendation

herdita mutiasari Logout

Hi!, See what we bring for you.

Select the algorithm and scroll down to see your recommendation list and don't forget to feedback us!

Select algorithm



3.3. Analisis K-Means Clustering

Untuk menguji apakah hasil analisis yang dilakukan pada aplikasi sudah benar maka dilakukan perhitungan manual dengan mengambil 15 data untuk dilakukan analisis.

Table 1. Nilai Centroid Awal

Nilai Centroid Awal		
	genre	rating
cluster 1	5	3
cluster 2	2	3

Table 2. Data Iterasi 1

No	Data film			Hasil Iterasi 1			
	Nama Film	Genre	Rating	C1	C2	Jarak Terpendek	Cluster
1	film 1	5	3	0.00	3.00	0.00	C1
2	film 2	2	3	3.00	0.00	0.00	C2
3	film 3	2	1	3.61	2.00	2.00	C2
4	film 4	5	4	1.00	3.16	1.00	C1
5	film 5	5	5	2.00	3.61	2.00	C1
6	film 6	4	5	2.24	2.83	2.24	C1
7	film 7	2	1	3.61	2.00	2.00	C2
8	film 8	3	1	2.83	2.24	2.24	C2
9	film 9	3	1	2.83	2.24	2.24	C2
10	film 10	3	2	2.24	1.41	1.41	C2
11	film 11	2	2	3.16	1.00	1.00	C2
12	film 12	4	3	1.00	2.00	1.00	C1
13	film 13	5	1	2.00	3.61	2.00	C1
14	film 14	4	3	1.00	2.00	1.00	C1
15	film 15	1	3	4.00	1.00	1.00	C2

Table 3. Nilai Centroid Baru

Nilai Centroid 2		
	genre	rating
cluster 1	5	3
cluster 2	2	2

Table 4. Data Iterasi 2

Data film			Hasil Iterasi 2			
Nama Film	Genre	Rating	C1	C2	Jarak Terpendek	Cluster
film 1	5	3	0.61	3.02	0.61	C1
film 2	2	3	2.61	1.27	1.27	C2
film 3	2	1	3.54	0.79	0.79	C2
film 4	5	4	0.71	3.55	0.71	C1
film 5	5	5	1.63	4.26	1.63	C1
film 6	4	5	1.67	3.69	1.67	C1
film 7	2	1	3.54	0.79	0.79	C2
film 8	3	1	2.89	1.06	1.06	C2
film 9	3	1	2.89	1.06	1.06	C2
film 10	3	2	2.12	0.79	0.79	C2
film 11	2	2	2.94	0.35	0.35	C2
film 12	4	3	0.71	2.15	0.71	C1
film 13	5	1	2.47	2.85	2.47	C1

film 14	4	3	0.71	2.15	0.71	C1
film 15	1	3	3.60	1.77	1.77	C2

Hasil : Data hasil antara iterasi 1 dan iterasi 2 identik (sama) maka proses iterasi akan dihentikan karena sudah menemukan kecocokan data sehingga menunjukkan hasil akhir perhitungan manual.

3.4. Pengujian Alfa

Pengujian alfa adalah pengujian yang difokuskan pada hasil akhir yang didapat, proses yang dilakukan dengan melakukan pengecekan pada setiap detail yang ada terhadap hasil akhir yang diinginkan apakah sudah sesuai atau belum. Pada scenario pengujian alfa, terdapat 7 fitur yang akan diuji. Pengujian alfa adalah pengujian dengan mengamati hasil eksekusi melalui data uji dan fungsionalitas dari aplikasi website.

Table 5. Pengujian Alfa

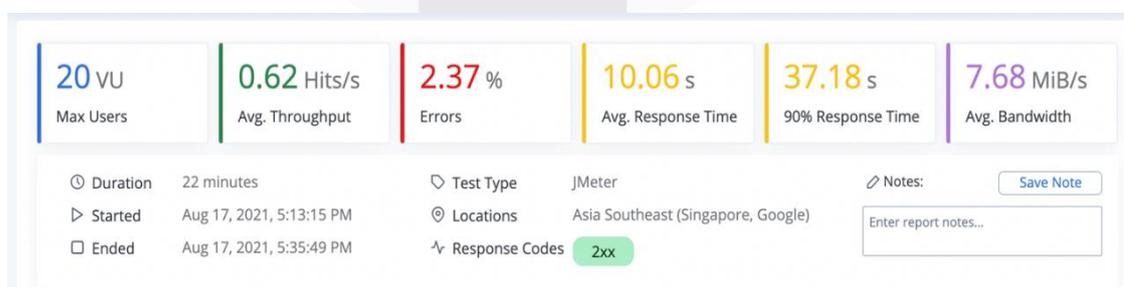
Fitur yang diuji	Detail Pengujian	Jenis Pengujian
Membuka aplikasi <i>website</i>	Membuka tampilan utama	Fungsional Test
Melakukan <i>Sign Up</i>	Melakukan pendaftaran akun	Fungsional Test
Melakukan <i>Login</i>	Melakukan masuk ke menu utama	Fungsional Test
Melakukan pengisian <i>GenreWeight</i>	Mengisi nilai untuk setiap genre yang dipilih	Fungsional Test
Membuka menu <i>Home</i>	Menampilkan halaman utama	Fungsional Test
Membuka menu Rekomendasi	Menampilkan hasil rekomendasi film	Fungsional Test
<i>Logout</i>	Keluar dari <i>website</i>	Fungsional Test

3.5. Pengujian Beta

Berdasarkan tabel hasil pengujian alfa yang dilakukan menunjukkan bahwa setiap komponen yang diujikan dapat memenuhi hasil yang diharapkan pada aplikasi yang dibangun. Dengan ini dapat disimpulkan bahwa fungsionalitas dari aplikasi yang dibangun dapat berfungsi baik. Dari hasil pengujian beta didapatkan hasil rata rata yang diperoleh adalah 4,32 dari range nilai 1 hingga 5 yang mana menunjukkan bahwa aplikasi website yang dibangun sudah berfungsi dengan baik dan dapat diterima oleh pengguna.

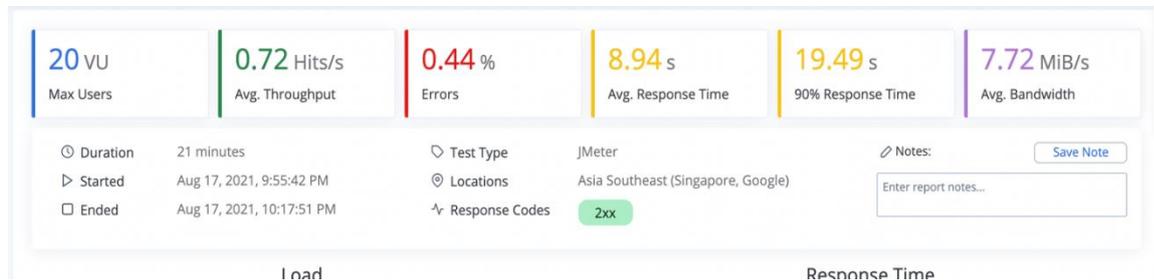
3.6. Pengujian JMeter

3.6.1. 3 Bobot Preferences Genre



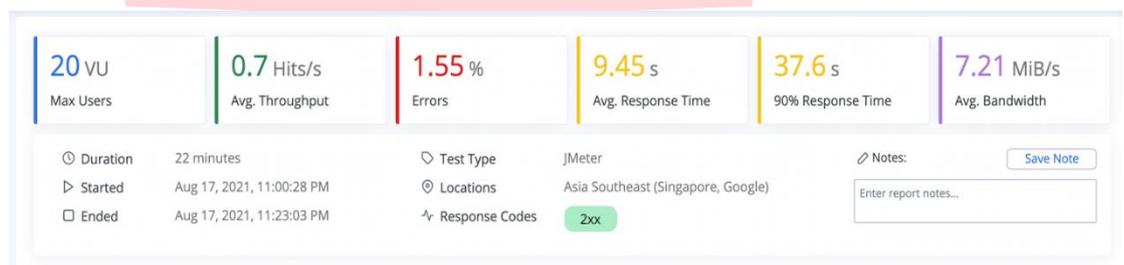
Pada gambar diatas merupakan hasil dari *stress test* dengan 20 *user*, terdapat 2,37% error yang terjadi.

3.6.2. 4 Bobot Preferences



Pada gambar diatas merupakan hasil dari stress test dengan 20 user, terdapat 0,44% error yang terjadi.

3.6.3. 5 Preferences



Pada gambar diatas merupakan hasil dari *stress test* dengan 20 user, terdapat 1,55% error yang terjadi.

4. Kesimpulan

4.1. Kesimpulan

Berdasarkan hasil perancangan, pengujian dan Analisa yang telah dilakukan maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Aplikasi ini mampu membantu pengguna untuk menampilkan list film yang sesuai dengan apa yang diinginkan oleh pengguna.
2. Analisis K-Means dapat mempermudah melakukan rekomendasi film sesuai dengan genre dan rating.
3. Aplikasi menggunakan 4 bobot preferences pada algoritma K-Means menghasilkan nilai error yang lebih kecil dengan nilai 0,44% dibandingkan menggunakan 3 atau 5 preferences.
4. Hasil dari genre yang tidak menggunakan bobot hasilnya sama dengan hasil yang menggunakan 3 preferences genre.

5.2. Saran

Saran untuk pengembangan aplikasi kedepannya adalah:

1. Bisa menambahkan menu rekomendasi berdasarkan tahun film untuk mempermudah pencarian film dan ditambahkan lagi trailer untuk filmnya.
2. Bisa menambahkan informasi terkait algoritma yang digunakan.
3. Terdapat algoritma metode pengklasifikasi yang lain sehingga pengguna dapat membandingkan efektifitas dari hasil klasifikasi untuk melakukan rekomendasi film.

Referensi:

- [1] W. Hanafiah, "PENINGKATAN KETERAMPILAN BERBICARA BAHASA INGGRIS MELALUI MEDIA FILM," vol. 16, no. 2, p. 149, 2019.
- [2] J. Maximillian, H. N. Palit and A. N. Tjondrowiguno, "Sistem Rekomendasi Film Menggunakan Integrated Kohonen K-Means Clustering".
- [3] A. Halim, H. Gohzali, D. M. Panjaitan and I. Maulana, "Sistem Rekomendasi Film menggunakan Bisecting K-Means dan Collaborative Filtering," *CITISEE 2017*, p. 37, 2017.
- [4] M.Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski," *Jurnal Informatika*, vol. 4, no. 1, 2019.
- [5] J. Fadhil and W. F. Mahmudy, "Pembuatan Sistem Rekomendasi Menggunakan Decision Tree dan Clustering," vol. 3, no. 1, 2007.
- [6] R. Ahuja, A. Solanki and A. Nayyar, "Movie Recommender System Using K-Means," 2019.
- [7] J. O. Ong, "IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING PRESIDENT UNIVERSITY," *Jurnal Ilmiah Teknik Industri*, vol. 12, no. 1, 2013.
- [8] S. Andayani, "Pembentukan cluster dalam Knowledge Discovery in Database dengan Algoritma K-Means," 2007.
- [9] Ediyanto, M. N. Mara and N. Satyahadewi, "PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTERING ANALYSIS," vol. 2, no. 2, 2013.