

## Sistem Rekomendasi untuk Maksimalisasi Industri Film dengan Metode Demographic Filtering dan Content Based Filtering

Puji Lestari\*<sup>1</sup>

<sup>1</sup>Ilmu Ekonomi, Fakultas Ekonomi dan Bisnis, Universitas Padjadjaran, Indonesia  
Email: <sup>1</sup>puji2002@mail.unpad.ac.id

### Abstrak

Saat ini terjadinya peningkatan *demand* terhadap layanan *streaming* terutama di industri perfilman, di tahun 2021 pendapatan yang dihasilkan dari langganan *video on demand* di Indonesia dapat mencapai USD 411 juta. Untuk memaksimalkan penggunaan layanan *streaming* ini diperlukan suatu sistem yang dapat menguntungkan produsen dan juga pengguna. Salah satu sistem tersebut adalah *movie recommendation system* yang dapat membantu pengguna dengan memberikan rekomendasi film terbaik dan sesuai dengan preferensi pengguna. Sistem ini juga dapat membantu perusahaan untuk menentukan target pasar dan memaksimalkan produksi film. Penelitian ini menggunakan dua metode dalam membuat rekomendasi sistem yaitu *Demographic Filtering* dan *Content Based Filtering*. Dengan menggunakan data yang berisi 3.315.117,054 *rating* yang berasal dari 4803 pengguna untuk 4801 film, sistem ini berhasil merekomendasikan 10 film dengan peringkat teratas bagi pengguna baru dengan menggunakan metode *Demographic Filtering* serta 10 film rekomendasi lainnya yang menyesuaikan dengan preferensi pengguna berdasarkan informasi sinopsis, genre, *credit*, dan *keyword* film dengan menggunakan metode *Content Based Filtering* sehingga rekomendasi yang diberikan lebih terpersonalisasi dengan preferensi pengguna. Kedua sistem dasar ini sangat efektif untuk digunakan karena dapat menguntungkan kedua pihak yaitu pengguna dan produsen film.

**Kata kunci:** *Content Based Filtering, Cosine Similarity, Demographic Filtering, Movie Recommendation System, TF-IDF*

### Abstract

Currently, there is an increasing demand for streaming services, especially in the film industry, in 2021 the revenue generated from video on demand subscriptions in Indonesia could reach USD 411 million. To maximize the use of this streaming service, a system is needed that can benefit producers and users. One such system is a movie recommendation system that can help users by providing the best movie recommendations and according to user preferences. This system can also help companies to determine the target market and maximize film production. This research uses two methods in making system recommendations, namely *Demographic Filtering* and *Content Based Filtering*. Using data containing 3,315,117,054 ratings from 4803 users for 4801 movies, this system successfully recommends 10 top-rated movies for new users using the *Demographic Filtering* method as well as 10 other recommended movies that adjust to user preferences based on synopsis, genre, credits, and movie keyword information using the *Content Based Filtering* method so that the recommendations provided are more personalized to user preferences. These two basic systems are very effective to use because they can benefit both parties, namely users and movie producers.

**Keywords:** *Content Based Filtering, Cosine Similarity, Demographic Filtering, Movie Recommendation System, TF-IDF*.

## 1. PENDAHULUAN

Diawali pada saat zaman Hindia Belanda pada tahun 1900 masyarakat Indonesia mulai mengenal film dengan sebutan gambar ideoep. Industri ini dimulai saat pembuatan film pertama di Hindia Belanda pada tahun 1926 dengan film yang berjudul *Loetoeng Kasaroeng* yang dibuat oleh L.Hoeveldorp dari NV Java Film Company yang dipimpin oleh G.Krugers dan F.Carli. Dimulai sejak saat itu industri perfilman di Indonesia mulai berkembang dan mengalami kondisi naik-turun, dimana sempat

mengalami titik terendah pada dekade 1990-an. Namun, mulai mengalami peningkatan ketika memasuki tahun 2000 [7], [11].

Berdasarkan studi yang dilakukan oleh konsultan dan peneliti dari Oxford Economics pada tahun 2010, yang menjelaskan bahwa perkembangan industri ini terus berlanjut dimana industri film di Indonesia mampu memberikan kontribusi pada Produk Domestik Bruto (PDB) nasional hingga 0,43% atau sebesar 2,98 juta dollar AS serta menciptakan lapangan kerja secara nasional sebesar 0,45% atau sebanyak 491.800 tenaga kerja [10].

Namun, saat terjadinya pandemi Covid-19 pada tahun 2020, industri perfilman menjadi salah satu sektor yang terdampak karena mulai terhentinya para pekerja seni untuk menghasilkan film-film hingga penutupan dan pembatasan bioskop untuk memutus penyebaran Covid-19. Akan tetapi, ditengah keterbatasan tersebut industri perfilman berhasil bangkit kembali dengan berbagai macam inovasi contohnya layanan *streaming* berbasis *platform digital* dengan *video on demand*. Inovasi ini berhasil karena berdasarkan data statistik, pendapatan yang dihasilkan dari langganan *video on demand* Indonesia dapat mencapai USD 411 juta di tahun 2021 dengan penetrasi pengguna sebesar 16% di tahun 2021 dan diperkirakan akan terus meningkat menjadi 20% di tahun 2025 [13].

Ditengah maraknya penggunaan layanan *film streaming* serta jumlah film yang terus meningkat maka dibutuhkan suatu sistem yang memadai agar penggunaan layanan *streaming* berbasis *platform digital* dengan *video on demand* ini dapat memberikan hasil yang maksimal baik untuk konsumen, produsen ataupun pemerintah. Salah satu sistem yang sangat berguna dalam memberikan layanan terbaik untuk pengguna adalah *movie recommendation system*. Dengan adanya sistem ini pengguna akan diberikan pengalaman menonton film yang baik karena akan diberikan rekomendasi film dengan *rating* terbaik dan sesuai dengan kategori film yang diinginkan pengguna [6].

Dengan kepuasan pengguna yang semakin meningkat terhadap layanan *film streaming* yang ada juga akan memberikan dampak yang baik pada produsen dan pemerintah karena *demand* akan film yang terus meningkat serta keuntungan yang didapatkan dari layanan tersebut. Tanpa adanya *movie recommendation system* pengguna akan kesulitan untuk menentukan film yang akan ditontonnya serta akan memakan waktu pengguna saat proses pencarian film belum lagi jika film yang sudah ditemukan tidak sesuai dengan selera atau keinginan pengguna yang pada akhirnya akan memakan waktu yang lebih lama lagi bagi pengguna untuk mencari film yang berkualitas dan sesuai dengan keinginannya. Jika hal itu terjadi maka kepuasan pengguna akan menurun dan juga akan berdampak pada *demand* serta keuntungan yang dihasilkan [15], [17].

Sistem rekomendasi film sendiri merupakan suatu *machine learning* yang membantu layanan *streaming* berbasis *platform digital* untuk memberikan informasi kepada pengguna mengenai film dengan *rating* terbaik serta merekomendasikan film favorit pengguna berdasarkan kepada preferensi dan perilaku pengguna. Sistem ini akan memberikan sejumlah list film yang sesuai dengan informasi yang diberikan oleh pengguna [14]. Sistem rekomendasi dapat berdasarkan kepada informasi *browsing history* pengguna, film apa yang sedang marak ditonton oleh pengguna dengan lokasi yang sama, serta *rating* dari pengguna mengenai film tersebut [18].

Terdapat beberapa manfaat yang diberikan oleh sistem rekomendasi, diantaranya yaitu menguntungkan pengguna dalam menemukan item yang mereka minati, membantu produsen untuk memberikan item atau produk ke pengguna yang tepat, mengidentifikasi produk yang paling relevan untuk setiap pengguna, konten yang dipersonalisasi, membantu *platform* untuk meningkatkan keterlibatan pengguna [14].

Penelitian empiris mengenai sistem rekomendasi diantaranya dilakukan oleh Girish Kumar P, Akhilesh M, yang melakukan penelitian pembuatan *product recommendation system* pada *e-commerce* Amazon dengan menggunakan pendekatan *machine learning* metode *Collaborative Filtering* dan *K-means*. Beberapa output yang dihasilkan yaitu informasi mengenai 10 data produk dengan jumlah *rating* tertinggi, memberikan rekomendasi barang serupa seperti yang sudah dibeli oleh konsumen dan membuat beberapa kluster produk yang memiliki karakteristik yang serupa. Dengan beberapa output yang dihasilkan oleh *machine learning* tersebut dapat membantu pelanggan dalam membuat keputusan membeli suatu produk dengan lebih efektif dan tepat [1].

Reddy S., *et al.* melakukan penelitian pada dataset *Movie Lens* dengan tujuan untuk membuat *movie recommendation system* dengan pendekatan *Content-Based Learning* berdasarkan korelasi antar jenis yang disukai oleh pengguna. Jadi, jika pengguna tersebut memberikan skor *rating* yang tinggi pada genre tersebut maka sistem akan memberikan rekomendasi film dengan genre yang serupa kepada pengguna [2].

Halim A., *et al.* membuat *movie recommendation system* dengan menggunakan metode *K-Means* dan *Collaborative Filtering* pada dataset *MovieLens* yang berisi 100.000 *rating* dari 668 pengguna. Implementasi untuk sistem ini adalah pembuatan *website* dengan menggunakan bahasa pemrograman PHP, HTML, dan CSS dengan basis datanya berupa MySQL. Peneliti juga melakukan pengujian model evaluasi antara kedua metode, ternyata kombinasi dari *bisecting K-Means* dan *User-based Collaborative filtering* memiliki nilai MAE yang lebih rendah dibandingkan *K-Means* dan *Item Based Collaborative filtering* yang artinya kombinasi *bisecting K-Means* dan *user-based* menghasilkan nilai rekomendasi yang lebih baik [3].

Fajriansyah M., *et all* melakukan penelitian mengenai algoritma *Content Based Filtering* untuk membuat *movie recommendation system* dengan mencari kemiripan bobot dari *term* pada *bag of words* hasil *pre-processing* sinopsis film dan judul film. Pembobotan dilakukan menggunakan metode TF-IDF yang telah dinormalisasi. Setelah itu dapat disimpulkan bahwa kemiripan bobot dari kueri, TF-IDF dan *cosine similarity* memiliki pengaruh terhadap hasil rekomendasi [4].

Gomez-uribe C.A., dan Hunt N. penelitian ini mendiskusikan beberapa algoritma yang digunakan Netflix untuk membuat *movie recommendation system*. Algoritma-algoritma tersebut adalah algoritma *Personalized Video Ranker (PVR)*, *Top-N Video Ranker*, *Trending Now Algorithm*, *Continue Watching Ranker*, *Video-Video Similarity*, *Page Generation: Row Selection and Ranking Algorithm*, *Evidence Algorithm*, dan *Search Algorithm* [5].

Untuk memaksimalkan pengalaman pengguna dalam layanan *streaming* serta membantu produsen untuk menemukan film yang paling diminati oleh pengguna maka penelitian ini dilakukan untuk membuat suatu sistem dasar yang dapat memberikan rekomendasi terbaik dengan berbagai informasi yang ada dalam dataset. Dataset yang digunakan adalah data sekunder yang berasal dari Full *MovieLens*, dataset tersebut maka akan diolah dengan menggunakan dua metode yaitu *Demographic Filtering* dan *Content Based Filtering* dengan harapan sistem tersebut akan memberikan rekomendasi film terbaik dan sesuai dengan karakter film yang disukai oleh para pengguna.

## 2. METODE PENELITIAN

Sistem rekomendasi adalah suatu sistem yang didesain untuk merekomendasikan suatu hal kepada pengguna berdasarkan beberapa faktor. Sistem rekomendasi akan memprediksikan produk yang paling banyak disukai oleh pengguna dan juga produk yang sesuai dengan selera dan keinginan pengguna tersebut [8], [9].

Dari berbagai macam informasi yang tersedia pada suatu *dataset*, sistem rekomendasi akan memfilter beberapa informasi penting dan faktor-faktor yang berkaitan dengan selera pengguna. Sistem rekomendasi akan mencoba untuk mencari kesamaan antara informasi yang diberikan oleh pengguna dengan informasi yang terdapat pada item tersebut untuk memberikan rekomendasi yang sesuai dengan keinginan pengguna [9].

Penulis menggunakan dua metode untuk membuat *movie recommendation system* yaitu *Content Based Filtering* dan *Demographic Filtering*. *Demographic Filtering* ditujukan untuk para pengguna baru dimana sistem akan memberikan 10 rekomendasi film berdasarkan *rating* tertinggi. Sedangkan *Content Based Filtering* ditujukan untuk para pengguna yang ingin menonton film yang sesuai dengan genre, produser ataupun atribut lainnya yang mereka sukai. Sehingga sistem akan merekomendasikan film yang berkorelasi dengan input yang diberikan oleh pengguna.. Dataset yang digunakan untuk membuat *machine learning ini* adalah *tmdb\_5000\_credits* dan *tmdb\_5000\_movies*.yang kemudian akan digabung berdasarkan id pelanggan yang ada pada kedua dataset [12].

Dataset tersebut berisi 3.315.117,054 *rating* berasal dari 4803 pengguna untuk 4801 film. Terdapat beberapa informasi yang ada dalam dataset yang diantaranya adalah informasi pemain, kru, *keywords*,

biaya, pendapatan, tanggal rilis, bahasa, perusahaan yang memproduksi film, negara, jumlah *vote* pada TMDB dan rata-rata *vote* pada TMDB [12].

## 2.1. Demographic Filtering

Rekomendasi sistem dengan metode *Demographic Filtering* mengklasifikasikan pengguna berdasarkan kepada informasi demografis dan merekomendasikan item-item yang sesuai dengan informasi tersebut. Dalam metode ini profil pengguna dibuat dengan mengklasifikasikan pengguna ke dalam deskripsi stereotip, yang merepresentasikan fitur-fitur kelas dari para pengguna [16],[19].

*Demographic Filtering* membuat kategori pengguna yang memiliki karakteristik demografis yang serupa [19]. Dalam kasus *movie recommendation system*, sistem akan memberikan rekomendasi umum untuk setiap pengguna berdasarkan kepada popularitas suatu film. Sistem akan memberikan rekomendasi yang sama kepada pengguna yang memiliki informasi demografis yang serupa [19]. Akan tetapi, karena setiap pengguna itu memiliki karakteristik yang berbeda maka pendekatan ini dianggap terlalu sederhana dan kurang mencerminkan personalisasi para pengguna. Ide dasar pada metode ini adalah bahwa film yang lebih populer dan banyak diakui secara kritis oleh para pengguna akan memiliki kemungkinan yang lebih tinggi untuk disukai oleh penonton secara rata-rata [12].

Algoritma yang dilakukan dalam metode ini adalah sebagai berikut [12]:

- Membuat matriks atau *rate movie*
- Menghitung skor untuk setiap film
- Menyortir film dan memberikan rekomendasi film terbaik untuk pengguna berdasarkan kepada jumlah skor rating tertinggi

Untuk menghitung skor *rating* dari setiap film sebenarnya dapat dilakukan dengan menggunakan nilai rata-rata dari setiap filmnya. Akan tetapi, cara ini dinilai kurang adil karena beberapa film memiliki *rating* 9 tetapi hanya memiliki 5 suara, sehingga tidak adil untuk film-film lain yang memiliki *rating* sedikit lebih rendah tetapi memiliki lebih banyak suara. Untuk menghitung nilai *rating* film yang lebih baik, IMDB sendiri sudah memiliki formulanya yang bernama *Weighted Rating (WR)*. Formula ini berdasarkan kepada jumlah suara, jumlah suara minimum, rata-rata suara, dan rata-rata *rating*. Berikut ini adalah formula *Weighted Rating* oleh IMDB [12]:

$$W = \frac{R*v+C*m}{v+m} \quad (1)$$

Keterangan:

$R$  : rata-rata film tersebut (*vote\_average*)

$v$  : jumlah *voting* film (*vote\_count*)

$C$  : rata-rata semua film

$m$  : syarat minimum *vote*

Karena untuk nilai  $R$  dan  $v$  sudah ada dalam dataset maka langkah selanjutnya adalah mencari nilai  $C$  dan  $m$  terlebih dahulu.

Code:

```
C= df2['vote_average'].mean()
```

Langkah selanjutnya adalah menentukan nilai  $m$  yang merupakan syarat minimum jumlah suara yang akan ditampilkan pada *chart recommendation system* nanti. Pada penelitian ini menggunakan persentil 90% sebagai *cutoff*. Dengan kata lain, untuk sebuah film dapat tampil di *chart* maka film tersebut setidaknya harus memiliki *vote* kurang lebih sebesar 90%.

Code:

```
m= df2['vote_count'].quantile(0.9)
```

Setelah mendapatkan nilai rata-rata semua film dan syarat minimum *vote* maka sistem akan membuat matriks yang berisi 10 rekomendasi film dengan menggunakan formula *Weighted Rating*.

## 2.2. Content Based Filtering

Pada sistem ini *machine learning* akan merekomendasikan suatu item berdasarkan kepada kemiripan atribut dari item yang disukai atau yang diinput oleh pengguna. *Content Based Filtering* juga merekomendasikan suatu item dengan menggunakan kata kunci dan atribut yang ditetapkan ke objek dalam *database* dan mencocokkannya dengan profil pengguna [9], [16].

Untuk kasus *movie recommendation system*, sistem ini akan menggunakan informasi dari metadata film tersebut, seperti genre, sutradara, deskripsi (sinopsis), aktor, dan informasi lainnya untuk memberikan informasi kepada pengguna. Ide umum dibalik sistem ini adalah jika pengguna menyukai suatu film tertentu maka pengguna tersebut juga akan menyukai film yang mirip atau memiliki karakteristik yang serupa dengan film tersebut [19].

Tidak seperti *Demographic Filtering*, pada metode ini setiap elemen atau informasi dalam film akan diperhitungkan dan dipertimbangkan sebagai dasar untuk membuat *Movie Recommendation System*. Dengan metode ini pengguna akan cenderung lebih menerima rekomendasi yang lebih sesuai dengan film favorit mereka [12].

### 2.2.1. Content Based Filtering Berdasarkan Deskripsi Film

Pada model ini sistem rekomendasi akan memberikan rekomendasi film kepada pengguna berdasarkan kemiripan deskripsi film atau sinopsis dari film yang pengguna tersebut sukai. Langkah pertama untuk menemukan kemiripan deskripsi antar film adalah dengan menghitung vektor *Term Frequency – Inverse Document Frequency (TF-IDF)* [12]. Algoritma *TF-IDF* adalah algoritma yang digunakan untuk menganalisis hubungan antara sebuah kata/kalimat dengan sekumpulan dokumen.

$$TF = \frac{\text{jumlah frekuensi kata terpilih}}{\text{jumlah kata dalam dokumen}} \quad (2)$$

$$IDF = \log \left( \frac{\text{jumlah dokumen}}{\text{jumlah frekuensi kata terpilih}} \right) \quad (3)$$

Setelah itu menghitung *TF-IDF* dengan cara sebagai berikut:

$$TF - IDF = TF * IDF \quad (4)$$

Untuk membuat *movie recommendation system* berdasarkan kesamaan deskripsi film maka langkah selanjutnya adalah menghitung kesamaan antar deskripsi film dengan menggunakan metode *Cosine Similarities* [12].

$$\text{Similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (5)$$

*Cosine Similarities* juga bisa diperoleh dengan menggunakan fungsi *linear\_kernel()* dari *sklearn library* pada Python. Setelah itu maka sistem akan memberikan 10 rekomendasi film setiap pengguna menginput 1 judul film yang mereka sukai [12].

### 2.2.2. Content Based Filtering Berdasarkan Aktor, Genre, Credit, dan keyword Film

Untuk meningkatkan kualitas sistem dan memberikan rekomendasi yang lebih sesuai dengan apa yang diinginkan oleh pengguna maka perlu ditambahkan informasi pada model. Penulis menambahkan informasi aktor, produser, genre dan juga *keyword* dari film yang diinput oleh pengguna. Untuk proses pembuatan *machine learning* sama seperti sebelumnya akan tetapi informasi yang diinput pada model lebih banyak dan metode yang digunakan untuk menghitung kemiripan deskripsi antar film adalah dengan *CountVectorizer()* dari *sklearn library* yang memiliki fungsi yang sama dengan *TF-IDF* yaitu untuk menghitung frekuensi kata dalam dokumen [12].



### 3. HASIL DAN PEMBAHASAN

Dari tiga pendekatan yang sudah dilakukan terdapat hasil 10 rekomendasi film yang berbeda-beda sesuai dengan pendekatan yang dilakukan.

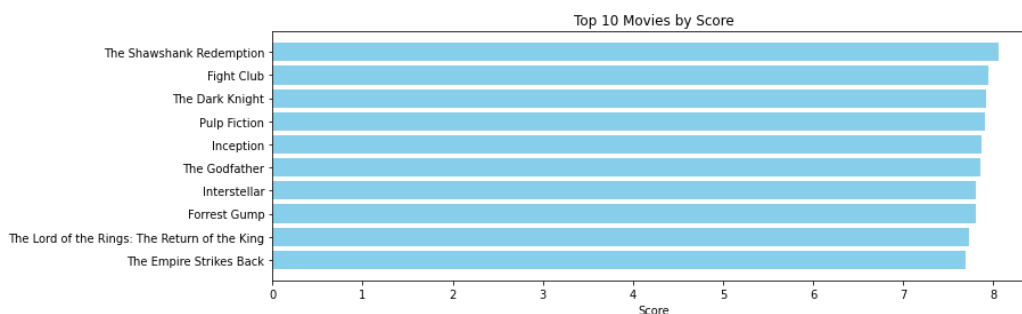
#### 3.1. Demographic Filtering

Metode ini menghasilkan nilai rata-rata atau *C score* dari dataset film tersebut sebesar 6.092171559442016. Maka, nilai rata-rata untuk semua film adalah sekitar 6 dari skala 10. Setelah mendapatkan nilai rata-rata maka sistem harus menentukan jumlah *vote* minimum agar film dapat tampil di *chart* rekomendasi yang akan dibuat. Setelah diproses maka jumlah *vote* minimal agar film tersebut dapat tampil di *chart* adalah sebesar 1.838,4. Setelah mendapatkan nilai *m* atau batas minimum dari jumlah *vote* yang dapat tampil pada *chart* maka sistem akan menyortir film dengan jumlah *vote* lebih dari 1.838,4 dan hasilnya adalah terdapat 481 film yang layak untuk tampil di *chart* dan direkomendasikan pada pelanggan. Dari 481 film tersebut maka akan diambil 10 film terbaik dengan menggunakan formula *Weighted Rating*. Berikut ini hasil *Weighted Rating* untuk 10 film:

Tabel 1. Hasil *Weighted Rating*

<i>Movie Title</i>	<i>Vote Count</i>	<i>Vote Average</i>	<i>Score WR</i>
The Shawshank Redemption	8205	8.5	8.1
Fight Club	9413	8.3	7.9
The Dark Knight	12002	8.2	7.9
Pulp Fiction	8428	8.3	7.9
Inception	13752	8.1	7.9
The Godfather	5893	8.4	7.9
Interstellar	10867	8.1	7.8
Forrest Gump	7927	8.2	7.8
The Lord of the Rings: The Return of the King	8064	8.1	7.7
The Empire Strikes Back	5879	8.2	7.7

Metode ini cocok untuk *platform streaming* yang tidak memiliki data *history* pengguna atau untuk pengguna baru. Dengan metode ini pengguna akan direkomendasikan film-film terbaik berdasarkan jumlah skor yang diberikan oleh pengguna lainnya yang sudah menonton film tersebut.



Gambar 1. Hasil Demographic *Filtering*

Sehingga, sistem dapat memberikan saran yang baik untuk pengguna yang belum memiliki pengalaman menonton film sebelumnya dan membantu pengguna untuk memilih film terbaik dari 4.801 film yang ada. Dari dataset ini 10 film dengan peringkat teratas berdasarkan jumlah skornya dapat dilihat pada Gambar 1. Dari *bar chart* tersebut dapat dilihat bahwa film dengan jumlah skor tertinggi adalah film “The Shawshank Redemption” dengan skor 8,1.

### 3.2. Content Based Filtering

Jika menggunakan metode *Demographic Filtering* pengguna diberikan rekomendasi 10 film berdasarkan jumlah skor tertinggi. Maka dengan metode *Content Based Filtering* sistem dapat memberikan rekomendasi film yang lebih terpersonalisasi dengan pengguna tersebut karena pengguna dapat menginput 1 film yang mereka sukai dan sistem akan memberikan 10 rekomendasi filmnya yang berkorelasi dengan film tersebut berdasarkan elemen-elemen atau informasi yang diinput pada saat *machine learning* untuk *movie recommendation* tersebut dibuat.

Pendekatan pertama yang kami lakukan untuk metode *Content Based Filtering* adalah dengan pendekatan *Content Based Filtering* berdasarkan deskripsi film. Dengan menghitung vektor *Term Frequency – Inverse Document Frequency (TF-IDF)* yang sudah dilakukan maka didapatkan hasil bahwa ada 20.978 kata yang digunakan untuk mendeskripsikan 4.803 film. Pada pendekatan ini juga sistem hanya akan memberikan 10 rekomendasi film yang memiliki tingkat kesamaan pada deskripsi atau sinopsis tertinggi dengan film yang diinput pengguna pada sistem ini tanpa mempertimbangkan elemen lainnya karena saat pembuatan *machine learning* ini yang diperhitungkan hanya kesamaan deskripsi antar film.

Model *machine learning*:

`get_recommendations('judul film')`

Contoh:

`get_recommendations('The Dark Knight Rises')`

Tabel 2. Hasil *Content Based Filtering* Berdasarkan Deskripsi Film

input	output film rekomendasi
The Dark Knight Rises	1. The Dark Knight
	2. Batman Forever
	3. Batman Returns
	4. Batman
	5. Batman: The Dark Knight Returns, Part 2
	6. Batman Begins
	7. Slow Burn
	8. Batman v Superman: Dawn of Justice
	9. JFK
	10. Batman & Robin

Sistem *machine learning* ini sudah berhasil untuk memberikan rekomendasi film yang memiliki deskripsi atau alur yang sama dengan yang pengguna inginkan. Pada contoh diatas jika pengguna menginput judul film “The Dark Knight Rises” maka sistem akan memberikan rekomendasi beberapa film yang memiliki cerita serupa atau serial dari film “The Dark Knight Rises” seperti “The Dark Knight”

Walaupun sistem tersebut sudah dapat memberikan rekomendasi film yang memiliki korelasi disetiap deskripsinya akan tetapi model ini bisa dikatakan belum memuaskan untuk dijadikan *movie recommendation system* karena informasi yang dimiliki sistem hanya informasi deskripsi dari film saja sehingga film yang direkomendasikan cenderung memiliki karakter cerita yang sama yang kemungkinan akan membuat pengguna merasa bosan dan cenderung kurang puas karena terus direkomendasikan film dengan cerita yang sama.

Sehingga untuk memperbaiki sistem yang ada maka dapat menggunakan pendekatan *Content Based Filtering* berdasarkan aktor, genre, *credit*, dan *keyword* film sehingga informasi yang dapat diolah oleh sistem menjadi lebih banyak dan dapat memberikan rekomendasi film yang lebih beragam tetapi tetap memiliki korelasi antar filmnya dengan film yang diinput pengguna pada sistem.

Model *machine learning*:

`get_recommendations('Judul film', cosine_sim2)`

Contoh:

`get_recommendations('The Dark Knight Rises', cosine_sim2)`

Tabel 3. *Content Based Filtering* Berdasarkan Aktor, Genre, *Credit*, dan *keyword* Film

input	output film rekomendasi
The Dark Knight Rises	1. The Dark Knight 2. Batman Begins 3. Amidst the Devil's Wings 4. The Prestige 5. Romeo Is Bleeding 6. Black November 7. Takers 8. Faster 9. Catwoman 10. Gangster Squad

Pada pendekatan ini *machine learning* berhasil untuk memberikan rekomendasi berdasarkan informasi aktor, genre, *credit*, dan *keyword* film yang ada pada sistem. Sehingga sistem bisa memberikan rekomendasi film yang memiliki korelasi dengan informasi yang diinput oleh pengguna dari berbagai elemen film.

Pada contoh diatas pengguna menginput judul yang sama dengan pendekatan *machine learning* yang sebelumnya. Akan tetapi film yang direkomendasikan berbeda oleh sistem berbeda. Pada *machine learning* yang sebelumnya rekomendasi film hanya terbatas pada deskripsi film tersebut sehingga cakupan film yang dapat direkomendasikan lebih sempit dibandingkan dengan cakupan film yang dapat direkomendasikan dengan menggunakan pendekatan *machine learning* ini.

Contohnya ketika pengguna menginput judul film “The Dark Knight Rises” maka sistem dapat merekomendasikan beberapa film yang memiliki korelasi pada setiap elemen film. Sistem akan merekomendasikan “The Dark Knight” karena film tersebut merupakan salah satu dari series “The Dark Knight Rises” karena memiliki deskripsi film yang serupa. Sistem juga merekomendasikan “The Prestige” karena baik film “The Dark Knight” ataupun “The Prestige” diproduksi oleh orang yang sama yaitu Christopher Nolan, walaupun kedua film memiliki alur atau deskripsi cerita yang berbeda.

#### 4. KESIMPULAN

*Movie recommendation system* dengan metode *Demographic Filtering* akan memberikan informasi mengenai film yang memiliki *rating* paling tinggi di *platform* sedangkan *Content Based Filtering* akan memberikan rekomendasi film yang fokus pada selera dan preferensi pengguna sehingga pada akhirnya para pengguna tidak perlu banyak menghabiskan waktu untuk mencari film dengan *rating* yang tinggi ataupun kumpulan film yang serupa dengan selera pengguna. Adanya informasi mengenai *rating* tertinggi di *platform* tersebut dapat mendorong perusahaan agar terus meningkatkan kualitas film yang dapat diproduksi agar film tersebut dapat masuk ke dalam *chart* film dengan jumlah *rating* tertinggi. Ketika film yang diproduksi masuk ke dalam *chart* tersebut maka akan lebih menarik minat pengguna untuk menonton film tersebut dan dapat meningkatkan pendapatan perusahaan pembuat film tersebut.

Selain memberikan informasi mengenai film dengan *rating* tertinggi, sistem ini juga dapat membantu pengguna untuk menemukan film yang sesuai dengan film yang mereka sukai sehingga dapat memberikan pengalaman menonton film yang lebih baik bagi pengguna. Untuk perusahaan juga sistem ini memiliki peran yang penting karena perusahaan menjadi dapat menentukan target pasar dari film yang mereka buat. Perusahaan dapat memprediksikan apakah film yang mereka buat dapat muncul direkomendasi pengguna atau tidak dengan melihat korelasinya dengan film-film lainnya. Sehingga target pasar yang dibuat perusahaan akan lebih tepat dan pengguna juga dapat menemukan film yang sesuai dengan yang mereka inginkan.



## DAFTAR PUSTAKA

- [1] Girish Kumar P, Akhilesh M. “IRJET- An Innovative Way Recommend Products in E-Commerce.” *Irjet* 8, no. 8, 2021.
- [2] Reddy, Srs, Sravani Nalluri, Subramanyam Kuniseti, S. Ashok, and B. Venkatesh. “Content-Based Movie Recommendation System Using Genre Correlation.” *Smart Innovation, Systems and Technologies*, no. 105, pp. 391–397, 2019.
- [3] Halim, Arwin, Hernawati Gohzali, Dita Maria Panjaitan, and Ilham Maulana. “Sistem Rekomendasi Filmmenggunakan Bisecting K-Means Dan Collaborative Filtering.” *Citisee* 4567789, no. 061, pp. 37–41, 2017.
- [4] Fajriansyah, Muhammad, Putra Pandu Adikara, and Agus Wahyu Widodo. “Sistem Rekomendasi Film Menggunakan Metode Content Based Filtering.” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 6, pp. 2188–2199, 2021. <http://ejournal.uajy.ac.id/20600/>.
- [5] Gomez-Uribe, Carlos A., and Neil Hunt. “The Netflix Recommender System: Algorithms, Business Value, and Innovation.” *ACM Transactions on Management Information Systems*, vol. 6, no. 4, 2015.
- [6] Agustian, Eggy Ryana. “Sistem Rekomendasi Film Menggunakan Metode Collaborative Filtering Dan K-Nearest Neighbors.” *Jurnal Infra*, vol. 3, no. 1, pp. 18–21, 2020. <http://publication.petra.ac.id/index.php/teknik-informatika/article/view/9800>.
- [7] Direktorat Sejarah, & Direktorat Jenderal Kebudayaan. (2017). Merayakan Film Nasional. Kementerian Pendidikan dan Kebudayaan Republik Indonesia. <https://repositori.kemdikbud.go.id/23892/1/Merayakan%20film%20nasional.pdf>.
- [8] Deng, Yue. *Recommender Systems Based on Graph Embedding Techniques: A Review*. *IEEE Access*. Vol. 10, 2022.
- [9] Dwivedi, R. (n.d.). *What are recommendation systems in machine learning?* Analytics Steps - A leading source of Technical & Financial content. <https://www.analyticssteps.com/blogs/what-are-recommendation-systems-machine-learning>.
- [10] BeritaSatu.com. (2012). Industri film Beri Kontribusi Pada Perkembangan Ekonomi. [beritasatu.com. https://www.beritasatu.com/news/53799/industri-film-beri-kontribusi-pada-perkembangan-ekonomi](https://www.beritasatu.com/news/53799/industri-film-beri-kontribusi-pada-perkembangan-ekonomi).
- [11] Haris Jauhari (ed.). *Layar Perak 90 Tahun Bioskop di Indonesia*, pp. 25
- [12] Ibtessama. (2020). *Getting started with a movie recommendation system*. Kaggle: Your Machine Learning and Data Science Community. [https://www.kaggle.com/code/ibtessama/getting-started-with-a-movie-recommendation-system/data?select=tmdb\\_5000\\_movies.csv](https://www.kaggle.com/code/ibtessama/getting-started-with-a-movie-recommendation-system/data?select=tmdb_5000_movies.csv)
- [13] KOMINFO, P. (2021). *Pemerintah Dorong Industri film Manfaatkan platform digital*. Website Resmi Kementerian Komunikasi dan Informatika RI. <https://www.kominfo.go.id/content/detail/37105/pemerintah-dorong-industri-film-manfaatkan-platform-digital/0/berita>
- [14] Muvi One. (2022). *Role of a movie recommender system in the streaming industry*. <https://www.muvi.com/blogs/movie-recommender-system.html>
- [15] Ryana Agustian, Eggy, and Eddy Prasetyo Nugroho. “Sistem Rekomendasi Film Menggunakan Metode Collaborative Filtering Dan K-Nearest Neighbors Film Recommendation System Using Collaborative Filtering Method and K-Nearest Neighbors”, vol. 3, no. 1, pp. 6765–6775, 2020, <https://ejournal.upi.edu/index.php/JATIKOM>
- [16] Ryngksai, Iateilang, and L Chameikho. “Recommender Systems : Types of Filtering”, vol. 3, no. 11, pp. 251–254, 2014.
- [17] Sandrya, Vincent, Wasino Wasino, and Desi Arisandi. “Sistem Rekomendasi Film Menggunakan Metode Multiple Attribute Utility Theory.” *Computatio : Journal of Computer*

*Science and Information Systems*, vol. 6, no. 1, pp. 19, 2022

- [18] Soham Das, S. (2022). *Build a movie recommendation system on your own*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2020/11/create-your-own-movie-movie-recommendation-system/>
- [19] Tommy. (2021). *How machine learning recommends movies for you*. Medium. <https://towardsdatascience.com/how-machine-learning-recommends-movies-for-you-6f418f26bcb2>