

## PEMBUATAN WEBSITE MOVIE DENGAN SISTEM RECOMMENDER FILM BERBASIS METODE COLLABORATIVE FILTERING

\*)Devi Dwi Purwanto, \*\*)Yosi Kristian  
\*)Sistem Informasi, \*\*)Teknik Informatika  
Sekolah Tinggi Teknik Surabaya  
<sup>1</sup>[devi@stts.edu](mailto:devi@stts.edu)  
<sup>2</sup>[yosi@stts.edu](mailto:yosi@stts.edu)

### ABSTRAK

Film adalah artefak budaya yang diciptakan oleh budaya tertentu, yang mencerminkan budaya, dan juga dapat mempengaruhi penontonnya. Film ini dianggap sebagai bentuk seni yang penting, sumber hiburan populer serta mengindoktrinasi orang. Unsur-unsur visual film bioskop memberikan kekuatan universal komunikasi. Pembuatan website Movie dengan Sistem Recommender Film Berbasis Metode Collaborative Filtering membantu orang Indonesia untuk mendapatkan informasi tentang film. Dengan website ini diharapkan tidak ada kesulitan untuk mendapatkan data film karena website ini dapat diakses dimanapun dengan menggunakan internet.

Pelanggan akan diberikan rekomendasi film berdasarkan film film yang disukai oleh pelanggan tersebut. Kemudian akan dicari film-film lain yang memiliki kemiripan dengan film-film yang disukai pelanggan. Film hasil pencarian siap dijadikan sebagai bahan rekomendasi untuk pelanggan yang bersangkutan. Sistem rekomendasi membantu pengguna untuk menemukan item yang ingin dibeli di situs e-commerce dengan menghasilkan prediksi daftar N atas item yang direkomendasikan untuk pengguna tertentu.

Pembuatan website Movie dengan Sistem Recommender Film Berbasis Metode Collaborative Filtering membantu user terutama penggemar film untuk memperoleh informasi mengenai dunia perfilman Indonesia baik baru maupun lama. Dengan menggunakan adjusted cosine similarity hasil yang didapat cukup baik dalam merekomendasikan film ke orang sesuai selera.

Kata kunci: Film, Adjusted - Cosine Similarity

### ABSTRACT

*Films are cultural artifacts created by specific cultures, which reflect those cultures, and, in turn, affect them. Film is considered to be an important art form, a source of popular entertainment and a indoctrinating — citizens. The visual elements of cinema give motion pictures a universal power of communication. Movie Website Development with Film Recommendation System Based on Collaborative Filtering is aimed to help user especially Indonesian people to get information about Indonesian Film. Through this website, user no longer meet difficulty to get information because this program can be accessed anywhere and anytime using internet connection.*

*Recommendations will be given to the customer based on movies that was liked by the customer. Then, other movies that have similarity with customer's favourite movie will be searched. Recommender Systems help users to find the items they would like to purchase at E-commerce sites by producing a predicted likeliness score or a list of top N recommended items for a given user .*

*Movie Website Development with Film Recommendation System Based on Collaborative Filtering helps user, especially fans of the film to obtain information about Indonesian cinema. By using the adjusted cosine similarity the result is good enough to recommend movie to user based on their taste.*

*Keywords : Film, Adjusted - Cosine Similarity*

## **1. PENDAHULUAN**

Kemunduran moral telah terjadi di masyarakat belakangan ini. Film merupakan salah satu bentuk harapan manusia tentang kehidupan. Film memiliki peranan yang cukup penting dalam kehidupan masyarakat. Melalui film, ada banyak hal yang dapat disampaikan. Dunia film yang gemerlap selalu mengundang keingintahuan masyarakat. Pesan – pesan komunikasi terwujud dalam cerita dan misi yang dibawa dalam bentuk drama, action, komedi, dan horor. Jenis – jenis film inilah yang dikemas oleh sutradara sesuai tendensi masing – masing. Awalnya, pada tahun 1900-an di Indonesia film dilahirkan sebagai tontonan umum karena semata-mata menjadi alternatif bisnis besar jasa hiburan di masa depan masyarakat kota. Dengan perkembangannya, film sempat dicap sebagai hiburan rendah masyarakat kota. Namun, sejarah membuktikan bahwa film mampu melakukan kelahiran kembali untuk kemudian mampu menembus seluruh lapisan masyarakat termasuk lapisan intelektual dan budayawan.

Film adalah sebuah barang komoditi dengan tingkat harga jual yang tinggi dan diminati oleh masyarakat. Belakangan ini, tingginya minat masyarakat akan film inilah yang mendasari dibuatnya Internet Movie Database. Banyaknya jumlah film yang ada di pasaran sekarang tidaklah memungkinkan apabila hanya dapat diakses melalui media cetak. Katalog perfilman memuat semua informasi mengenai film yang dimaksud, seperti judul, sutradara, aktor, genre, sinopsis dan rating.

Internet merupakan sarana informasi yang sangat berkembang saat ini. Berbagai macam hal dapat dilakukan dengan menggunakan internet. Transaksi melalui internet merupakan salah satu manfaat yang penting dalam penggunaan internet, dimana seorang pembeli dapat menghibur diri, mendapatkan semua informasi dari segala bidang kehidupan tanpa harus pergi ke suatu tempat tujuan untuk melihat barang-barang yang dicari.

Pada IMDb, masyarakat diberikan kemudahan untuk dapat mencari film berdasarkan kata kunci yang diinputkan oleh pengguna dengan mudahnya tanpa harus mencari secara manual layaknya di media cetak. Selain itu, masyarakat dapat melihat berbagai film baik terbaru maupun lama sewaktu – waktu dan dimanapun dengan mengakses internet. Pengguna dapat membagikan informasi dan komentar mengenai film yang bersangkutan dengan pengguna lain. Internet Movie Database ini dapat juga digunakan sebagai wadah berkumpulnya penggemar film untuk saling berbagi informasi dengan orang lain. Dalam website juga terdapat sistem rekomendasi yang akan diberikan untuk membantu pengguna memilih film sesuai dengan minat pengguna. Setiap pengguna dipersilahkan untuk memberikan rating pada film yang sudah ditontonnya. Dari rating

inilah, akan ditentukan film mana yang direkomendasikan sistem untuk dapat ditonton oleh pengguna.

## 2. LANDASAN TEORI

Definisi Film menurut UU 8/1992, 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, elektronik, dan/atau lainnya.

### 2.1 Collaborative Filtering

*Collaborative filtering* adalah sistem rekomendasi yang menentukan profil dari selera pengguna berdasarkan informasi yang didapat dari pengalaman pengguna lain terutama pengguna yang memiliki ciri khas yang hampir sama dengan pengguna aktif. Dengan alasan tersebut, maka *collaborative filtering* menjadi suatu teknologi rekomendasi yang menjanjikan hasil terbaik dalam merekomendasikan sesuatu saat ini. *Collaborative filtering* memberikan rekomendasi berdasarkan kolaborasi antar pengguna. Dengan kata lain *collaborative filtering* digunakan untuk memberikan rekomendasi bagi pengguna berdasarkan informasi yang didapat dari pengguna lain.

*Collaborative filtering* dapat diartikan sebagai setiap algoritma yang dapat melakukan pemilahan terhadap informasi yang diinginkan pengguna berdasarkan profil dari pengguna. Sekelompok pengguna yang memiliki kesamaan profil berkemungkinan memiliki tingkat keingintahuan yang sama (*similar interest*). Untuk pengguna, informasi yang ada dapat difilter dipersempit atau diperlebar tergantung dari tingkat kesamaan terhadap sekelompok pengguna.

Aplikasi yang menggunakan *collaborative filtering* biasanya menggunakan data dalam ukuran yang besar. *Collaborative filtering* merupakan suatu proses dalam memberikan rekomendasi kepada seorang pengguna berdasarkan kolaborasi antar pengguna. Dengan kata lain *Collaborative filtering* digunakan untuk memberikan rekomendasi bagi pengguna berdasarkan informasi yang didapat dari pengguna lain. Dengan demikian *Collaborative filtering* dapat disebut sebagai *social filtering*.

### 2.2 Adjusted Cosine Similarity

Pengimplementasian Collaborative Filtering dengan metode apapun, pada dasarnya dapat diklasifikasikan menjadi 3 (tiga) tahapan penting, yaitu data preparation, pencarian pola selera user dan penyusunan rekomendasi. Ketiga tahapan tersebut merupakan satu kesatuan solid yang berkaitan satu sama lain. Begitu pula halnya dengan metode yang digunakan dalam penelitian ini. Tahap pertama, data preparation, tahap berikutnya, yaitu pencarian pola selera pelanggan yaitu pembentukan tabel yang nantinya akan membentuk tabel akhir yang diistilahkan dengan nama tabel similarity. Tahap terakhir, pemberian rekomendasi.

#### 2.2.1. Data Preparation

Pada fase data preparation, pengguna memberikan rating ke film yang pernah dilihat. Rating yang diberikan berkisar antara 1 sampai dengan 10. Nilai 1 diberikan dengan asumsi bahwa film tersebut tidak patut untuk dilihat, dari nilai 1 akan beranjak

ke 2, 3 dan seterusnya sampai dengan nilai 10 yang diasumsikan bahwa film tersebut patut untuk ditonton. Tabel 1 berisikan 10 film beserta rating dan berapa banyak orang yang rating. Field dari rating didapat dari rata – rata setiap rating yang diberikan oleh setiap user. Berikut akan diberikan contoh pemberian rating dari 5 user sampai dengan ditemukan nilai similarity dari user 1 ke user lain.

Tabel 1. Contoh Tabel Film

Judul	Rating	Jumlah orang Rating
Love	3	1
Pocong vs KuntilAnak	6	4
Claudia	5.5	2
Love Is Cinta	2	2
Fiksi	5.5	2
Ayat – Ayat Cinta	3	3
I Love You Om	6	1
Sang Dewi	2	1
Tentang Cinta	4	1
Lewat Djam Malam	3	1

Field rating pada contoh tabel film (tabel 2) berasal dari jumlah rating setiap user yang pernah merating film tersebut dibagi dengan jumlah orang yang merating film bersangkutan. Field jumlah orang rating berisikan jumlah orang yang pernah merating film tersebut. Setiap user yang merating film tertentu disimpan datanya ke suatu tabel yaitu tabel rating. Dalam tabel rating disebutkan setiap user merating film apa saja dan berapa ratingnya seperti pada contoh tabel rating pada tabel 2.

Tabel 2. Contoh Tabel Rating

	User I	User II	User III	User IV	User V
Love	3				
Pocong vs KuntilAnak	5	7	7		5
Claudia	7	4			
Love Is Cinta	1	3			
Fiksi		8		3	
Ayat – Ayat Cinta		3	3		3
I Love You Om				6	
Sang Dewi				2	
Tentang Cinta			4		
Lewat Djam Malam					3

Sebagai contoh pada tabel 2 Pocong vs Kuntilanak memiliki rating 6 dan jumlah rating 4. Jumlah rating 4 didapat dari 4 orang yang merating yaitu User I, User II, User III, dan User V yang dapat dilihat pada tabel 3. Field rating dari tabel 2 didapat dari rating User I terhadap Pocong vs Kuntilanak ditambah dengan rating User II terhadap Pocong vs Kuntilanak ditambah dengan rating User III terhadap Pocong vs Kuntilanak ditambah dengan rating UserIV terhadap Pocong vs Kuntilanak. Jadi perhitungannya adalah  $(5+7+7+5)/4=6$ .

### 2.2.2. Pencarian Pola

Setelah melewati tahap preparation (persiapan), data yang didapat siap diolah untuk menemukan CI. Langkah berikutnya adalah menentukan nilai kemiripan / *similarity* suatu user dengan user lainnya. Nilai tersebut dikenal dengan istilah *Commonality Index* (CI). Proses ini dapat dikatakan sebagai proses pencarian pola selera pelanggan yang terdaftar dan pernah memberikan rating pada suatu film. Semakin tinggi nilai CI suatu pasangan user menandakan bahwa kedua user tersebut memiliki kemiripan dalam hal selera film. Algoritma yang dipakai dalam penelitian ini adalah *adjusted cosine similarity*.

$$\text{Sim}(i,j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}}$$

Keterangan :  
 R(u,i) = rating dari user u dengan item i.  
 R(u) = rata-rata dari user u

Dalam algoritma diatas, dibandingkan antara similarity dari film 1 ke film lainnya, sedangkan pada Penelitian ini similarity di modifikasi menjadi similarity antara user. Algoritma akan menghitung kedekatan user 1 dengan user lainnya.

Setelah semua data yang berada di tabel film dan tabel rating di hitung menggunakan algoritma *adjusted cosine similarity*. Hasil perhitungan similarity antar user dimasukkan ke tabel similarity seperti pada tabel 3.4. Semua similarity dihitung apabila user sama – sama merating film yang sama. Film yang tidak dirating oleh salah satu dari user tidak dihitung. Sebagai contoh perhitungan dari hasil tabel similarity diambil dari user I dan user II sebagai berikut:

$$\frac{(7-6)(7-6) + (0-5.5)(4-5.5) + (0-2)(3-2) + (0-5.5)(8-5.5) + (3-3)(3-3) + (4-4)(0-4)}{\sqrt{(7-6)^2 + (0-5.5)^2 + (0-2)^2 + (0-5.5)^2 + (3-3)^2 + (4-4)^2} \times \sqrt{(7-6)^2 + (4-5.5)^2 + (3-2)^2 + (8-5.5)^2 + (3-3)^2 + (0-4)^2}}$$

Pada contoh perhitungan terdapat nilai 0, apabila terdapat nilai 0 berarti tidak dihitung karena rating 0 dan tidak merating berbeda. Dalam perhitungan tersebut perhitungan (0-5.5)(4-5.5), (0-2)(3-2), (0-5.5)(8-5.5), (4-4)(0-4) tidak dianggap. Begitu pula dengan penyebut, nilai (0-5.5), (0-2), (0-5.5), (4-4), (4-5.5), (3-2), (8-5.5) serta (0-4) juga tidak dihitung. Setelah nilai yang tidak dihitung hilang akan tersisa

$$\frac{(7-6)(7-6) + (3-3)(3-3)}{\sqrt{(7-6)^2 + (3-3)^2} \times \sqrt{(7-6)^2 + (3-3)^2}} = 1$$

Tabel 3. Contoh Tabel Similarity

	User I	User II	User III	User IV	User V
User I	1	1	-1	0	-1
User II		1	-1	-1	-1
User III			1	0	1
User IV				1	0
User V					1

### 2.2.3. Penyusunan Rekomendasi

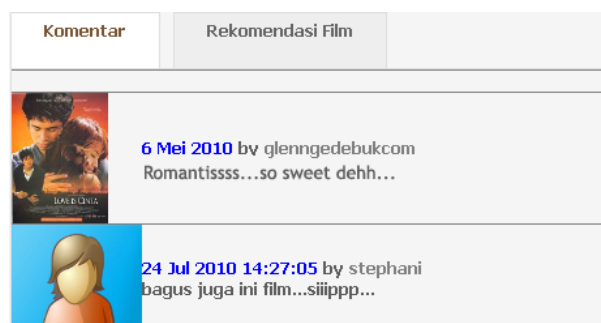
Pada tahap ini tabel similarity sudah terbentuk seperti yang terlihat pada tabel 3. Bila dilihat pada tabel 3 dapat disimpulkan bahwa tingkat similarity pasangan User I dengan User II, User II dengan User IV, User III dengan User I, User IV dengan User II serta User V dengan User IV memiliki similarity yang paling tinggi. Semakin tinggi hasil similarity, tingkat kesamaan film antar user juga akan tinggi.

Setelah mendapatkan hasil similarity dan mengetahui user yang memiliki tingkat similarity paling tinggi, 3 user yang memiliki CI tertinggi diambil untuk dicari film yang pernah dirating dan dipakai sebagai rekomendasi film. Berikut akan dijelaskan dengan menggunakan salah satu contoh. User I memiliki similarity paling tinggi dengan user II yaitu 1. Film yang pernah dirating oleh user II akan menjadi rekomendasi untuk user I. Karena dalam penelitian ini diambil 3 user yang memiliki tingkat kesamaan yang tertinggi, maka User I menerima rekomendasi dari user II, user III, user IV. Film yang pernah dirating oleh user II, user III, dan user IV akan direkomendasikan.

## 3. DESAIN INTERFACE

Pada halaman utama akan menampilkan menu login pada bagian kiri. Selain dapat melakukan login sesuai username dan password pada bagian tersebut pemakai juga dapat melakukan pendaftaran dengan menekan link daftar yang disediakan. Dibawah login terdapat fasilitas gambar aktor, yang akan menampilkan gambar aktor secara acak dari katalog film. Pada bagian tengah akan menampilkan 5 film secara acak dan juga menampilkan 6 film baru yang memiliki rating tertinggi.

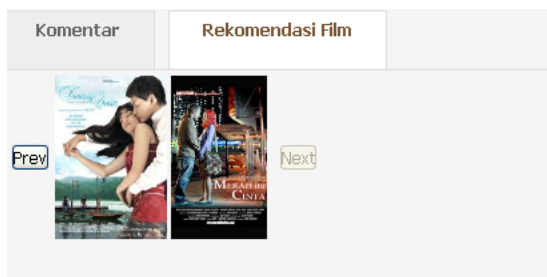
Halaman detail film akan menampilkan data film secara lengkap. Data tersebut meliputi judul, kategori, pemain, produksi, produser, penulis, sutradara, website untuk melihat trailer, sinopsis, warna, tagline, tanggal tayang, bahasa, lama tayang, rating film, nilai, komentar, sinopsis dan trailer dari suatu film dan penghargaan yang pernah diraih. Selain menampilkan data data film, pada halaman detail film pelanggan dapat memberikan penilaian rating dan komentar untuk film tersebut. Jangkauan untuk nilai rating adalah 1 sampai dengan 10. Asumsi nilai 10 adalah nilai yang diberikan apabila film dianggap sangat bagus. Rekomendasi akan menampilkan sejumlah film yang dirating oleh 3 orang yang memiliki tingkat kedekatan paling tinggi dengan pengguna. Pengguna hanya dapat memberikan nilai rating, komentar dan mendapatkan rekomendasi hanya bila pengguna telah melakukan registrasi serta login sebagai anggota situs IdMDb.



Gambar 1. Halaman Pemberian Komentar

Komentar yang dimasukkan oleh pengguna tidak langsung secara otomatis ditampilkan. Apabila komentar tidak mengandung kata – kata kotor dan layak dibaca

oleh khalayak umum, maka komentar yang dimasukkan pengguna akan otomatis tampil.



Gambar 2. Halaman Fitur Rekomendasi Film

Nama : Grace

Film yang dirating	Rating	Rekomendasi yang diberikan	Selera (suka/tidak/belum ditonton)
Anak Setan	4	Skandal Cinta Babi Ngepet	Tidak
Love	5	Mas Suka Mas Ukin Aja	suka
Jelangkung 3	3	Denias	suka
Mas Suka Mas Ukin Aja	7	Jelangkung 3	suka
Hantu Aborsi	4	Bukan Malin Kundang	tidak
		Kamulah satu-satunya	suka
		Planet Mars	Belum ditonton
		Kawin Kontrak Lagi	Suka
		Anak Setan	Suka
		Oh Baby	Suka
		Aborsi	Suka
		Love	Suka
		Merah itu Cinta	Belum ditonton
		Tapi Bukan Aku	Belum ditonton
		Bukan Cinta Biasa	Suka
		SMS	Suka
		Perjaka Terakhir	Tidak

Total Film Disukai : 11

Total Film Tidak Disukai : 3

Total Film Belum Ditonton : 3

Rata Rata Keakuratan :  $11 / 14 * 100\% = 78,57\%$

Gambar 3. Hasil Uji Coba dan Perhitungannya

#### 4. UJI COBA

Untuk uji coba terhadap hasil rekomendasi film, awalnya sukarelawan akan diberi pertanyaan film apakah yang disukai dan pernah ditonton oleh sukarelawan tersebut. Kemudian akan dicari rekomendasi berdasarkan film yang disukai sukarelawan tersebut, dan sistem rekomendasi akan memberikan sejumlah film yang disukai dengan rekomendasi berbunyi "pengguna yang merating film A akan merating film ini", dimana kata "film ini" mengacu pada film yang disukai dan pernah ditonton oleh sukarelawan

tadi. Dari sejumlah film yang dihasilkan oleh sistem rekomendasi akan ditanyakan kembali kepada sukarelawan, apakah sukarelawan tersebut juga pernah menonton film yang dijadikan bahan rekomendasi. Untuk detail dari uji coba dapat dilihat pada gambar 3.

## 5. KESIMPULAN

Berdasarkan pembahasan yang ada pada bagian sebelumnya maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Algoritma similarity yang sebenarnya merupakan similarity dari film ke film. Pada Penelitian ini diganti dengan similarity user ke user. Dengan menggunakan similarity user ke user hasil yang didapat cukup baik.
2. Kelemahan dari adjusted cosine similarity ini karena hasil yang didapat bergantung pada selera user.
3. Untuk memberikan suatu rekomendasi kepada seorang user, user harus memberikan rating yang tinggi terhadap beberapa film. Hal ini dilakukan agar sistem dapat mengetahui selera pelanggan untuk melakukan proses penyusunan rekomendasi.
4. Permasalahan yang ada pada proses rekomendasi adalah masalah waktu yang diperlukan untuk membentuk tabel similarity.

## 6. DAFTAR PUSTAKA

- Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl. *Item-based Collaborative Filtering Recommendation Algorithm*. <http://www10.org/cdrom/papers/519/>. May 2001.
- D. Goldberg, D. Nichols, B. M. Oki dan D. Terry. *Using collaborative Filtering to Weave an Information Tapestry*. Desember 1992.
- Dody Soegiarto. *Analisis Error-Tolerant Fascicles Untuk Collaborative Filtering*, <http://www.hansmichael.com/default.asp?cat=taF>. Penelitian Sekolah Tinggi Teknik Surabaya. 2004.
- F. Heylighen. *Collaborative Filterin.*, <http://pespmc1.vub.ac.be/COLLFILT.html>. January 2001.
- Marguerie, Fabrice. 2008. LINQ in Action - THE LINQ book. <http://linqinaction.net>. Diakses tanggal 13 Maret 2011.
- Siregar, Edison. *Membangun Aplikasi Berbasis Web dengan ASP.NET*. Gramedia. 2002.
- Telkom. Collaborative Filtering. 2009. [http://www.itelkom.ac.id/library/index.php?view=article&catid=20%3Ainformatika&id=507%3Acollaborative-filtering&option=com\\_content&Itemid=15](http://www.itelkom.ac.id/library/index.php?view=article&catid=20%3Ainformatika&id=507%3Acollaborative-filtering&option=com_content&Itemid=15). Diakses tanggal 1 September 2010.