Sistem Rekomendasi Film dengan Menggunakan *Sentiment Analysis* dan *Collaborative Filtering*

Muhamad Agus Fikry[1], Septiyawan Rosetya Wardhana[2],* dan *Rinci Kembang Hapsari[3]*

[1,2,3]Jurusan Teknik Informatika, Institut Teknologi Adhi tama Surabaya

**Penulis Korespondensi : rossywardhana@itats.ac.id*

ABSTRACT

Along with the rapid development of digital today, the era is increasingly advanced, and various kinds of files can be accessed from the internet. Likewise, the films we often watch on TV can now be easily accessed from the internet. Many film enthusiasts are sometimes still confused when they want to watch a movie. Referring to the description, in this study, a system is built that can provide film recommendations. Collaborative filtering is a method that is often used in terms of recommendations, and sentiment analysis is used to determine the sentiment pattern of the user. Python 3 programming language calculates the processes in the system to be created. From the references and methods, this study aims to build a recommendation system using Collaborative Filtering and Sentiment Analysis methods for film reviews. Testing was performed 5 times, and the results could not meet the expectations for recommendations because the average recommendation value was still 32%.

Article History

Received : 17-04-2025
Revised : 28-04-2025
Accepted : 01-05-2025

Key words

Collaborative Filtering,
Film,
Python,
Sentiment Analysis,
System Recommendation.

ABSTRAK

Seiring pesatnya perkembangan digital pada saat ini, zaman semakin maju berbagai macam file bisa diakses dari internet. Begitu juga dengan film yang sering kita tonton dari TV sekarang bisa diakses dari internet dengan mudah. Banyak peminat film yang kadang masih bingung ketika ingin menonton film. Mengacu pada uraian tersebut, dalam penelitian ini, membangun sebuah sistem yang dapat memberikan rekomendasi film. *Collaborative filtering* adalah metode yang sering digunakan dalam hal rekomendasi dan *Sentiment analysis* digunakan untuk menentukan pola sentiment dari user serta menggunakan bahasa pemrograman python 3 untuk menghitung proses-proses pada sistem yang akan dibuat. Dari acuan dan juga metode tersebut tujuan dari penelitian ini adalah membangun sistem rekomendasi menggunakan metode *Collaborative Filtering* dan *Sentiment Analysis* terhadap ulasan pada film. Pengujian dilakukan sebanyak 5 kali uji coba, dimana hasil belum bisa memenuhi harapan dalam merekomendasikan, karena rata-rata nilai rekomendasi masih 32%.

PENDAHULUAN

Di era perkembangan digital yang sedang berkembang sangat pesat ini banyak teknologi canggih baik itu media, ataupun infomarsi dan salah satunya adalah film. Di mana film yang dahulu biasa kita tonton dari televisi ataupun bioskop sekarang bisa ditonton melalui internet dengan cara streaming maupun download terlebih dahulu. Di era ini terdapat lonjakan konten hiburan digital yang belum pernah terjadi sebelumnya, pengguna sering menghadapi tantangan dalam menjelajahi katalog yang luas untuk menemukan film yang sesuai dengan preferensi mereka.

Seiring dengan terus berkembangnya perpustakaan platform streaming, permintaan akan sistem rekomendasi yang dipersonalisasi pun meningkat, yang bertujuan untuk meningkatkan kepuasan pengguna dengan memberikan saran yang relevan. Metode rekomendasi tradisional, meskipun bermanfaat, sering kali hanya mengandalkan penilaian pengguna yang eksplisit, yang mungkin tidak mencakup seluruh spektrum sentimen penonton. Untuk mengatasi keterbatasan ini, kemajuan terkini memadukan analisis sentimen dari ulasan yang dibuat pengguna dengan teknik penyaringan kolaboratif, yang memungkinkan pemahaman yang lebih bernuansa tentang preferensi pengguna dan respons emosional terhadap film[1]. Dengan menggabungkan pendekatan ini, sistem dapat menyempurnakan rekomendasi berdasarkan tidak hanya pada riwayat tontonan dan faktor demografi, tetapi juga pada sentimen yang diungkapkan dalam ulasan tekstual, sehingga meningkatkan akurasi dan keterlibatan pengguna [2]. Penggabungan ini merupakan langkah signifikan menuju sistem rekomendasi film yang dinamis dan sadar emosi.

Dalam beberapa tahun terakhir, preferensi untuk pengiriman konten yang dipersonalisasi telah meningkat dengan pesatnya pertumbuhan platform hiburan digital, yang membutuhkan sistem rekomendasi yang lebih canggih. Sistem rekomendasi film memanfaatkan berbagai algoritme, termasuk penyaringan kolaboratif, penyaringan berbasis konten, dan pendekatan hibrida, untuk menyarankan film yang disesuaikan dengan preferensi pengguna individu. Penyaringan kolaboratif memanfaatkan data interaksi pengguna seperti peringkat dan riwayat tontonan untuk mengidentifikasi kesamaan di antara perilaku pengguna, sementara analisis sentimen menggunakan teknik pemrosesan bahasa alami untuk menafsirkan ulasan yang dibuat pengguna, menambahkan lapisan bernuansa dengan menangkap respons emosional terhadap film. Dengan menggabungkan metode ini, sistem dapat memberikan rekomendasi yang lebih akurat dan bermakna yang melampaui sekadar peringkat numerik. Integrasi ini mengatasi tantangan yang terkait dengan kelangkaan data dan masalah cold-start, yang pada akhirnya meningkatkan kepuasan pengguna melalui penyelarasan yang tepat dengan selera dan sentimen pribadi. Akibatnya, model canggih yang menggabungkan analisis sentimen bersama penyaringan kolaboratif mencontohkan arah masa depan untuk teknologi rekomendasi film[3],[1].

Pada penelitian sebelumnya telah menggunakan pendekatan prediksi *Collaborative Filtering* dengan mengimplementasikan *deep learning* berdasarkan teknologi *Neural Collaborative Filtering* pada dataset MovieLens. Evaluasi model dilakukan menggunakan metrik skor regresi *Root Mean Square Error* (RMSE). Hasil pada pengujian model menunjukkan hasil terbaik dengan nilai rata-rata loss value sebesar 0,1356 pada fase train dan sebesar 0,8898 pada fase val, dengan learning rate dan batch size memperoleh kinerja terbaik ketika learning rate bernilai 0,001 dan batch size dengan nilai 1024 [4].

Pada penelitian yang lain metode *Collaborative Filtering* dengan jenis *item-based* telah bisa memberikan rekomendasi buku kepada pengguna. Dimana untuk menentukan kemiripan antar buku menggunakan *cosine similarity.* Dalam sistem rekomendasi yang dihasilkan tidak ditemukan kesalahan dalam menentukan rekomendasi buku [5]. Sehingga dalam penelitian ini menggunakan metode *analysis sentiment* dan *colaborative filtering* untuk memberikan rekomendasi film.

TINJAUAN PUSTAKA

Collaborative Filtering

Collaborative filtering adalah teknik yang digunakan untuk sistem rekomendasi dalam menyarankan suatu item dari satu user ke user lainnya yang memiliki kemiripan [6]. *Collaborative filtering* (CF) adalah metode yang digunakan untuk membuat prediksi dengan memperkirakan kesamaan ketertarikan antara user pada suatu sistem dan dipresentasikan dalam bentuk rating [7]. Secara umum, ada 2 proses yang dilakukan dalam CF, yaitu:

1. Mencari user lain yang mempunyai kemiripan pola rating dengan user target (user yang akan diberikan prediksi).
2. Menggunakan nilai rating dari user lain yang didapat dari langkah 1 di atas untuk menghitung prediksi bagi user aktif.

Sentiment Analysis

Sentiment Analysis atau disebut juga opinion mining adalah bidang studi yang digunakan untuk menganalisis pendapat orang, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan sebagainya.

TFIDF

Algoritma *Term Frequency Invers Document Frequency* atau sering disebut TFIDF adalah algoritma yang digunakan untuk menghitung nilai bobot pada suatu kata maupun dokumen. Sebelum melakukan perhitungan TFIDF biasanya akan dilakukan tahap *preprocessing* seperti yang

diuraikan diatas. TF-IDF (*Term Frequency Invers Document Frequency*) merupakan metode numerik yang statis yang menggambarkan pentingnya sebuah kata dalam sebuah korpus [8]. Pada algoritma TF-IDF digunakan rumus untuk menghitung bobot (w) masing-masing film dengan rumus yaitu :

$$w_{ij} = tf_{ij} \; x \; \log\left(\frac{N}{df_i}\right) \quad (1)$$

Dimana :

w_{ij} = bobot kata-*i* pada dokumen ke-*j*
tf_{ij} = banyak kata-*i* pada dokumen ke-*j*
N = total dokumen
df_i = banyak dokumen yang mengandung kata-*i*

Python

Python adalah satu dari berbagai macam bahasa pemrograman. Sama seperti bahasa yang digunakan manusia, pada komputer juga terdapat berbagai macam bahasa pemrograman seperti Java, LISP, PHP, dan Perl serta C dan yang lainnya. Dalam setiap bahasa pemrograman setidaknya ada satu hal yang bagus dari bahasa tersebut, misalnya menulis program yang mudah adalah point untuk bahasa pemrograman Java dan untuk mengakses database dan menyatukannya ke dalam halaman web adalah spesialisasi dari bahasa pemrograman PHP. Tapi, semua bahasa pemrograman ini pada intinya memiliki konsep yang sangat mirip. sebagian besar memiliki data dalam variabel dan fungsi (procedure, method) untuk melakukan berbagai macam hal pada suatu data. Beberapa bahasa bahkan menggabungkan data dan fungsi ke dalam bundel yang disebut objek, dan yang lainnya seperti LISP memungkinkan Anda memperlakukan fungsi seperti variabel, dan sebaliknya. Python adalah bahasa pemrograman yang kuat dan elegan yang mudah dibaca dan dipahami. Ini ditunjukkan sebagian besar fitur python banyak untuk umum.

METODE

Gambaran Umum

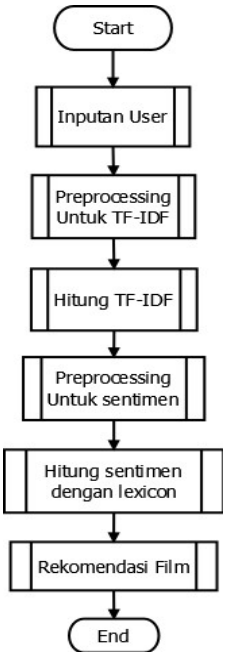
Pada penelitian memiliki tujuan yaitu memudahkan pengguna dalam mencari film dengan memberikan rekomedasi sesuai label dari pengguna lain, dimana label pada film akan dihitung dari berapa banyak sentimen positif dari user review yang terdapat pada film tersebut. pengambilan dataset secara manual pada situs film imdb.com dengan jumlah 330 data. Dataset nantinya akan berjumlah 330 data yang berisi sinopsis film dan user review pada masing-masing film.

Rancangan Sistem

Perancangan sistem adalah menuangkan suatu keadaan (input, proses dan output) yang ada pada sistem ke dalam suatu model tertentu. Alur flowchart dari rancangan sistem yang dibangun ditunjukkan pada Gambar 1.

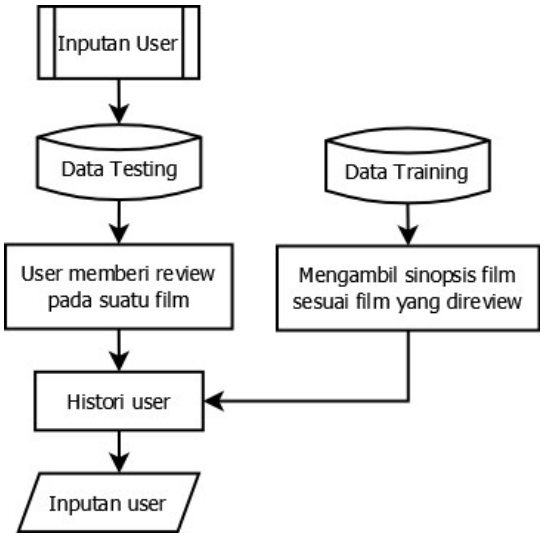
Pada Gambar 1 terdapat beberapa proses yang akan terdapat pada rancangan sistem pada penelitian. Proses-proses tersebut terdiri dari Inputan User, Preprocessing untuk TF-IDF, Menghitung TF-IDF, Preprocessing untuk sentimen, Menghitung sentimen menggunakan Lexicon opinion, sampai dengan Hasil rekomendasi film

ISSN: 2774-4345, DOI: 10.31284/j.kernel.2024.v5i2.7315

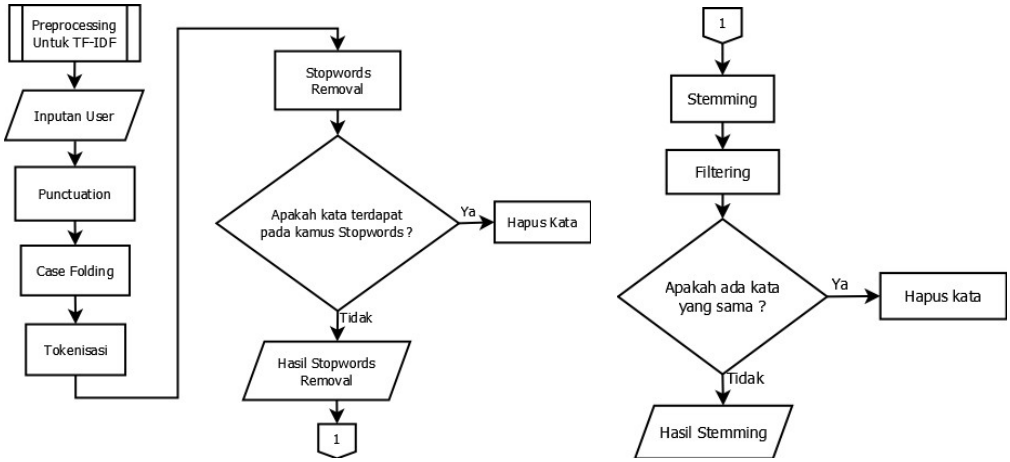


Gambar 1 Flowchart Rancangan Sistem

. Detail input sistem yang dilakukan oelh user ditunjukkan pada Gambar 2. Pertama inputan user akan mengambil user review dari Data Testing untuk memberi review pada suatu film dan akan disimpan pada Histori user. Setelah user memberi review pada suatu film secara otomatis Histori user akan mengambil sinopsis film yang sesuai dari Data Training dan akan disimpan.

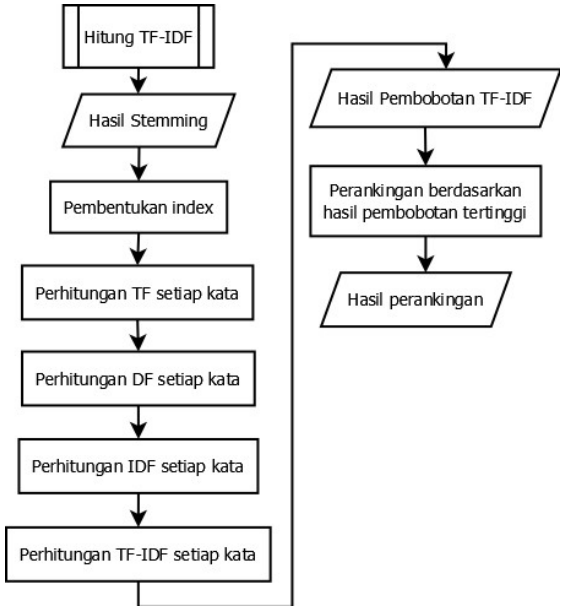


Gambar 2. Alur Input data oleh user



Gambar 3. Alur pre-proses untuk TF-IDF

Dalam pre-proses dilakukan beberapa sub proses, antara lain : *punctuation, case folding*, tokenisasi, dan *stopword removal*. Dimana alur adri prosesnya ditunjukkan pada Gambar 3. Selanjutnya adalah alur proses menghitung TF-IDF yang ditunjukkan pada Gambar 4. Hasil *preprocessing* dari langkah sebelumnya akan dibentuk index tiap kata (*Term*). Kemudian menghitung *Term Frequency* pada setiap film dan menghitung *Document Frequency* pada setiap film.



Gambar 4. Alur perhitungan TF-IDF

Selanjutnya menghitung *Inverse Document Frequency* pada setiap film, dilanjutkan dengan menghitung *TF-IDF* pada setiap film dan akan mendapatkan hasil bobot kata. Hasilkan dijadikan sebagai acuan perankingan berdasarkan bobot tertinggi dan akan diambil beberapa teratas untuk diproses ke langkah analisis sentiment.

HASIL DAN PEMBAHASAN

Hitung TFIDF

Pada pengujian disini peniliti menghitung bobot dari keseluruhan synopsis yang akan dijadikan data ranking untuk direkomendasikan kepada user lain. Dilakukan dengan menggunakan 4 data synopsis pada film dan sudah dilakukan tahapa preprocessing.

Tabel 1. Data Bobot Kata

	W = TF*IDF				
Term	**Histori user**	**D1**	**D2**	**D3**	**D4**
Big	0,30103	0,30103	0	0	0
Friend	0,124939	0,124939	0	0,124939	0
Kingdom	0	0	0	0,30103	0,30103
New	0,60206	0,60206	0	0	0
World	0,124939	0,124939	0,124939	0	0
Discov	0	0	0,30103	0,30103	0
Total	1,152968	1,152968	0,425969	0,726999	0,30103

Tabel 2. Data Bobot Kata

Judul Film	**Bobot Kata**
Toy Story 4	1,152968
Frozen II	0,726999
How to Train Your Dragon: The Hidden World	0,425969
The Lion King	0,30103

Hitung Sentiment Analysis

Pada pengujian sentiment analysis, peneliti memberikan contoh hasil perhitungan sentiment analysis dengan menggunakan lexicon-based. berikut rumus menghitung sentiment :

$$S = \frac{\sum P}{(\sum P + \sum Neg)} \quad (2)$$

Tabel 3. Data Sentiment Film

Toy Story 4				
User review	KK Positif	KK Negatif	Score	Sentimen
Review 1	3	2	0,6	Positif
Review 2	1	0	1	Positif
Review 3	1	2	0,3333	Negatif
Review 4	1	0	1	Positif
User Testing	2	0	1	Positif

KESIMPULAN

Dari pengujian yang sudah dilakukan pada bab 4, dapat ditarik kesimpulan yaitu dari 5 kali uji coba yang sudah dilakukan oleh tiap user masih belum bisa memenuhi harapan dalam rekomendasi, dikarenakan ada 4 user yang memiliki presentase dibawah 50% dan nilai rata-rata prosentaseny adalah 35%.

DAFTAR PUSTAKA

[1] A. A. Patel, V. V. Ubhe, S. G. Vaidya, K. Deskmukh, and V. Visal, “Integrating Sentiment Analysis and Machine Learning for Improved Film Recommendation Systems,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. XI, 2024.

[2] Ishika Naskar, & Niju P Joseph, "Impementation of Movie Recommendation System Using Hybrid Filtering Methods and Sentimen Analysis of Movie Reviews", I*EEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE*), Februari 2024.

[3] Pavitha N., Vithika Pungliya, Ankur Raut, et al., "Movie Recommendation and Sentiment Analysis Using Machine Learning", *Global Transitions Proceedings*, April 2022.

[4] Ni’mah Khoiriyah Ayyiyah, Retno Kusumaningrum, Rismiyati, " Film Recommender System Menggunakan Metode Neural Collaborative Filtering", *Jurnal Teknologi Informasi dan Ilmu Komputer,* Vol. 10 No. 3, 2023.

[5] Pharadya Ajeng Swari Sukmawati, Lely Hiryanto, & Viny Christanti Mawardi, "Implementasi Metode Collaborative Filtering Based untuk Sistem Rekomendasi Buku Fiksi", *Jurnal Ilmu Komputer dan Sistem Informasi,* Vol 11 No 2, 2023.

[6] Lisniati Dzumiroh & Ristu Saptono, " Penerapan Metode *Collaborative Filtering* Menggunakan Rating Implisit pada Sistem Perekomendasi Pemilihan Film di Rental VCD", *JURNAL ITSMART,* Vol 1. No 2. Desember 2012

[7] K Obajha, N N K Sari, & V H Pranatawijaya, "Implementasi Metode Collaborative Filtering pada Aplikasi Rekomendasi Hotel dan Wisma di Kota Palangka Raya Berbasis Website", *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi,* Vol. 3 No. 2, Desember 2023

[8] Leskovec, J., Rajaraman, A., & Ullman, J. D. (2014). Mining of Massive Datasets. *Mining of Massive Datasets*. https://doi.org/10.1017/cbo9781139924801