

Sistem Rekomendasi Film Berbasis Web App

Afalah Fakhri Andika Dirgantara
Data Science Faculty of Informatics
Telkom University
Bandung, Indonesia
adirgantara21@gmail.com

Afeef Radithya Rashid
Data Science Faculty of Informatics
Telkom University
Bandung, Indonesia
afeefradithya@gmail.com

Aflah Iram Athena
Data Science Faculty of Informatics
Telkom University
Bandung, Indonesia
aflahiram26@gmail.com

Abstract—Sistem rekomendasi film berbasis web app telah menjadi alat yang esensial untuk membantu pengguna menemukan film yang sesuai dengan preferensi mereka. Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan sistem rekomendasi film berbasis web menggunakan teknik Content-Based Filtering. Dataset yang digunakan berasal dari GitHub, yang berisi Top 1000 IMDB movie. Sistem ini mengumpulkan data dari preferensi pengguna dan informasi film seperti genre, sutradara, dan aktor untuk menghasilkan rekomendasi yang dipersonalisasi. Implementasi menggunakan teknologi web sederhana seperti Streamlit dan menggunakan bahasa pemrograman Python. Evaluasi sistem dilakukan melalui uji coba dengan sejumlah pengguna, menunjukkan bahwa sistem mampu memberikan rekomendasi yang relevan dan meningkatkan kepuasan pengguna dalam menemukan film baru. Dengan hasil ini, sistem rekomendasi film berbasis web dapat diintegrasikan lebih luas ke platform hiburan digital untuk meningkatkan pengalaman pengguna.

Keywords—Sistem Rekomendasi, Film, IMDB, Content-Based Filtering.

I. PENDAHULUAN

Dengan pesatnya perkembangan teknologi informasi dan komunikasi, platform hiburan digital telah menjadi bagian penting dalam kehidupan sehari-hari. Platform seperti Netflix, Amazon Prime, dan Disney+ menawarkan berbagai pilihan film dan serial TV yang sangat banyak sehingga pengguna sering kali kesulitan menemukan konten yang sesuai dengan preferensi mereka. Untuk mengatasi masalah ini, sistem rekomendasi telah dikembangkan dan diimplementasikan dalam berbagai platform untuk membantu pengguna menemukan konten yang sesuai dengan minat mereka.

Sistem rekomendasi menggunakan berbagai teknik untuk menganalisis data pengguna dan konten untuk memberikan rekomendasi yang dipersonalisasi. Salah satu teknik yang populer adalah Content-Based Filtering, yang menghasilkan rekomendasi berdasarkan kesamaan antara item yang disukai pengguna sebelumnya dan item yang tersedia. Teknik ini sangat efektif dalam menyarankan film yang memiliki genre, sutradara, atau aktor yang sama dengan film yang telah disukai pengguna.

Penelitian ini fokus pada pengembangan sistem rekomendasi film berbasis web app menggunakan teknik Content-Based Filtering. Dataset yang digunakan dalam penelitian ini berasal dari GitHub, yang berisi informasi mengenai Top 1000 IMDB movie. Dengan memanfaatkan dataset ini, sistem dapat memberikan rekomendasi film yang relevan berdasarkan preferensi pengguna.

II. DASAR TEORI

A. Content-Based Filtering

Content-Based Filtering (CBF) adalah teknik rekomendasi yang memberikan saran berdasarkan karakteristik atau fitur

dari item yang telah dinikmati pengguna sebelumnya. Dalam rekomendasi film, fitur-fitur ini mencakup genre, sutradara, aktor, dan deskripsi. Sistem CBF membangun profil pengguna berdasarkan fitur-fitur dari film yang disukai pengguna, lalu menggunakan metrik kesamaan, seperti cosine similarity, untuk menemukan film serupa.

CBF menghasilkan rekomendasi yang sangat personal dan menghindari masalah cold start untuk item baru, namun memiliki kelemahan seperti keterbatasan dalam menemukan item yang benar-benar baru dan kompleksitas dalam ekstraksi fitur. Implementasi CBF melibatkan pengumpulan data, ekstraksi fitur, pembangunan profil pengguna, perhitungan kesamaan, dan penyusunan rekomendasi.

Evaluasi sistem dilakukan dengan mengukur metrik seperti precision dan recall, serta mendapatkan umpan balik pengguna untuk menilai kepuasan dan efektivitas sistem. Memahami prinsip-prinsip ini penting untuk mengembangkan sistem rekomendasi berbasis CBF yang efektif.

B. Cosine Similarity

Cosine similarity adalah metrik yang digunakan untuk mengukur kesamaan antara dua vektor dalam ruang berdimensi tinggi dengan menghitung kosinus dari sudut di antara mereka. Nilai cosine similarity berkisar antara -1 hingga 1, di mana 1 berarti vektor-vektor tersebut identik, 0 berarti tidak ada kesamaan, dan -1 berarti vektor-vektor tersebut berlawanan arah. Cosine similarity sering digunakan dalam sistem rekomendasi untuk menentukan seberapa mirip dua item atau profil pengguna.

$$\text{Cosine Similarity} = (A, B) = \frac{A \cdot B}{\|A\| \|B\|}$$

III. METODE PENELITIAN

Dataset ini diperoleh dari github untuk digunakan sebagai bahan analisis. Dataset ini memiliki 16 kolom serta 1000 baris data. Namun setelah melakukan pemilihan fitur dan pembersihan data, kami hanya menggunakan 5 kolom dan 843 data. Berikut merupakan sebagian dari data yang kami gunakan untuk penelitian ini:

TABLE I. DATASET

Series_ Title	Genre	Overview	IMDB_ Rating	Meta_ Score
The Shaws hank Redem ption	Drama	Two imprisoned men bond..	9.3	80.0

The Godfather	Crime, Drama	An organized crime dynasty's aging...	9.2	100.0
The Dark Knight	Action, Crime, Drama	When the menace known...	9.0	84.0
...	...			
From Here to Eternity	Drama, Romance, War	In Hawaii in 1941, a private...	7.6	85.0
Lifeboat	Drama, War	Several survivors of a torpedoed...	7.6	78.0
The 39 Steps	Crime, Mystery, Thriller	A man in London tries to help a counter-espion...	7.6	93.0

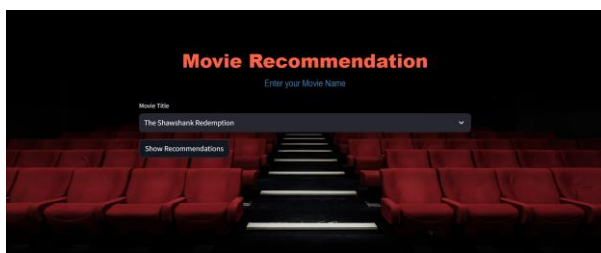
Selanjutnya untuk fitur `IMDB_Rating` dan `Meta_Score` akan dilakukan normalisasi terlebih dahulu supaya kedua fitur tersebut dapat digunakan. Langkah selanjutnya sebelum membangun model yang digunakan adalah membagi dataset secara random menjadi dataset latih dan uji dengan perbandingan 80:20. Langkah terakhir sebelum membangun model adalah melakukan vektorisasi teks untuk kolom Genre dan Overview.

Model dibangun menggunakan metode *embedding* terhadap vektorisasi teks nya sebagai itu lalu disatukan dengan kolom `IMDB_Rating` dan `Meta_Score` untuk input tambahannya. Penentuan kemiripan antar data nanti akan diukur menggunakan metode *cosine similarity*.

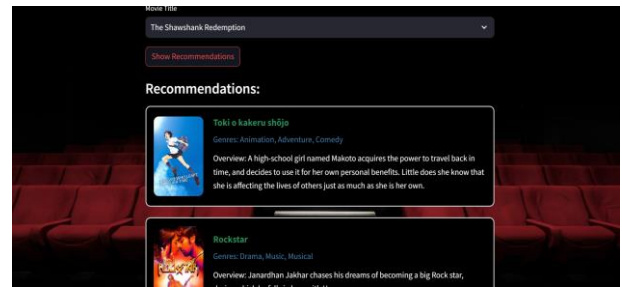
IV. HASIL

Dari sistem rekomendasi film yang telah dibuat, frontend yang digunakan untuk mengimplementasikan model yang telah dibuat adalah Streamlit. Streamlit adalah sebuah framework open-source yang digunakan untuk membuat aplikasi web interaktif dengan menggunakan kode Python. Berikut merupakan tampilan web app yang telah dibuat.

TAMPILAN WEB APP



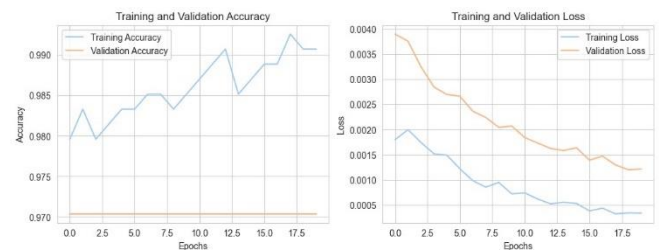
HASIL REKOMENDASI FILM



Pada web app ini, pengguna dapat memasukkan judul film yang diminati, seperti contohnya 'The Shawshank Redemption', dan mendapatkan rekomendasi film berdasarkan kemiripan dengan film tersebut.

Selain itu, pengguna dapat mengeksplorasi detail dari film-film yang direkomendasikan, seperti 'genre' dan 'overview'. Ini memberikan pengalaman yang lebih dalam dan membantu pengguna dalam membuat keputusan yang lebih baik terkait film-film yang ingin mereka tonton selanjutnya.

HASIL TRAINING



Pada grafik akurasi di sebelah kiri, terlihat bahwa akurasi pelatihan (ditunjukkan oleh garis biru) meningkat secara bertahap dari sekitar 0.980 pada awal pelatihan hingga mencapai sekitar 0.990 pada akhir pelatihan. Sementara itu, akurasi validasi (ditunjukkan oleh garis oranye) tetap stabil di sekitar 0.970 sepanjang proses pelatihan, tanpa menunjukkan peningkatan yang signifikan.

V. KESIMPULAN

Hasil dari pelatihan model ini dapat digunakan dengan baik pada *webapp* kami untuk menampilkan rekomendasi film berdasarkan tipe, deskripsi, skor IMDB, dan skor Meta. Yang dapat dikembangkan dari penelitian ini adalah penggunaan dataset yang lebih besar dan beragam. Dataset yang kami gunakan untuk penelitian ini merupakan dataset yang memuat film-film yang sudah cukup lama dan dikarenakan dataset ini berisikan 1000 film teratas pada situs IMDB, yang mana situs ini melakukan pengurutan berdasarkan voting, maka banyak dari film pada dataset kami terbilang bias dan sulit dikenali.

REFERENCES

- [1] K. Raj, A. A. Das, A. Guha, P. Sharma and M. K. S, "Movie Recommendation System," p. 1028, 2019.Fortunato, S. (2010). Community detection in graphs. Physics Reports, 486(3-5), 75-174.I.
- [2] J. Han, M. Kamber and J. Pei, Data Mining: Concepts and Techniques, 2012.
- [3] S. Bapodara, "Movie Recommendation System + Web App," 2023. [Online]. Available: <https://www.kaggle.com/code/sagarbapodara/movie->

recommendation-system-web-app/notebook. [Accessed 13 March 2024].