

PENERAPAN ALGORITMA CONTENT BASED FILTERING UNTUK SISTEM REKOMENDASI FILM DAN LAYANAN *STREAMING* FILM

Ade Kurnia Putra

ABSTRAK

Banyaknya pilihan film dan layanan *streaming* seringkali menyulitkan pengguna dalam pengambilan keputusan. Penelitian ini bertujuan untuk mengatasi kesulitan tersebut dengan merancang sebuah sistem rekomendasi film dan layanan *streaming* film berbasis *content-based filtering*. Data film diperoleh dengan teknik *scraping* dari *API The Movie Database* dengan jumlah 15.813 data film. Kemudian data tersebut melalui tahapan *preprocessing* dengan tahapan *case folding*, *cleaning*, *tokenizing*, dan *stemming*. Selanjutnya model dibuat menggunakan teknik *CountVectorizer* untuk mengubah fitur konten tekstual film menjadi vektor dan *Cosine Similarity* untuk mengukur kemiripannya. Model yang sudah dibuat akan di evaluasi menggunakan metrik evaluasi *MAP@K* yang mendapatkan skor terbesar pada *MAP@1* dengan skor 0.7577. Kemudian model diintegrasikan pada *website* yang dibuat menggunakan framework *ReactJs* sebagai *frontend* dan *Flask* sebagai *backend*. Fungsionalitas *website* diuji dengan *black box testing* dengan hasil berhasil pada semua skenario. Selanjutnya penilaian pengguna diuji menggunakan *user acceptance testing* dengan metode *beta testing* yang mendapatkan skor akhir 4,58 dari skala 5. Hasil ini membuktikan bahwa pendekatan *content-based filtering* dengan *CountVectorizer* dan *Cosine Similarity* terbukti efektif dan cocok untuk membangun sistem rekomendasi film dan layanan *streaming* film yang personal dan memuaskan bagi pengguna.

Kata Kunci: *Content based filtering*, *CountVectorizer*, *Cosine Similarity*, Film, Layanan *Streaming* Film

IMPLEMENTATION OF CONTENT BASED FILTERING ALGORITHM FOR MOVIE RECOMMENDATION SYSTEM AND MOVIE STREAMING SERVICES

Ade Kurnia Putra

ABSTRACT

The abundance of movies and streaming services often hinders user decision-making. This research aims to overcome these difficulties by designing a movie recommendation system and movie streaming services based on content-based filtering. Movie data is obtained by scraping from The Movie Database API with a total of 15,813 data. The data goes through preprocessing stages of case folding, cleaning, tokenizing, and stemming. Model is created using CountVectorizer technique to convert the film's textual content into vectors and Cosine Similarity to measure the similarity. The model is evaluated using MAP@K metrics which gets the largest score on MAP@1 with a score of 0,7577. The model is integrated on the website created using ReactJs framework as the frontend and Flask as the backend. The website functionality was tested with black box testing with successful results in all scenarios. Furthermore, user assessment was tested using user acceptance testing (UAT) with beta testing method which received a final score of 4.58 out of 5. These results prove that content-based filtering approach with CountVectorizer and Cosine Similarity is effective and suitable for building movie recommendation systems and movie streaming services that are personalized and satisfying for users.

Keywords: Content based filtering, CountVectorizer, Cosine Similarity, Movie, Movie Streaming Services