

ABSTRAK

Perkembangan layanan *streaming* menyebabkan jumlah film yang tersedia semakin banyak sehingga pengguna sering mengalami kesulitan dalam menemukan film yang sesuai dengan preferensinya. Penelitian ini bertujuan mengembangkan sistem rekomendasi film berbasis web menggunakan metode Content-Based Filtering (CBF) dengan memanfaatkan teknik Natural Language Processing (NLP) untuk mengolah data teks. Sistem dikembangkan menggunakan IMDb Movies Dataset yang diperoleh dari Kaggle dengan memanfaatkan atribut overview (sinopsis), genre, pemeran (*crew*), dan judul film sebagai dasar dalam menghitung tingkat kemiripan, sedangkan tahun rilis digunakan sebagai proses penyaringan (*filtering*) hasil rekomendasi sesuai preferensi pengguna. Tahapan pengolahan data meliputi *text preprocessing*, representasi teks menggunakan N-Gram (unigram dan bigram), pembobotan Term Frequency–Inverse Document Frequency (TF-IDF), perhitungan kemiripan menggunakan Cosine Similarity, serta penggabungan nilai kemiripan menggunakan metode average similarity. Sistem diimplementasikan menggunakan bahasa pemrograman Python sebagai mesin rekomendasi dan framework Laravel sebagai antarmuka pengguna. Evaluasi sistem dilakukan menggunakan metrik Precision@10 pada beberapa skenario pengujian dan menghasilkan nilai rata-rata sebesar 0,78, yang menunjukkan bahwa sistem mampu memberikan rekomendasi film yang relevan sesuai preferensi pengguna. Selain itu, evaluasi pengguna menggunakan pendekatan Technology Acceptance Model (TAM) memperoleh nilai rata-rata keseluruhan sebesar 3,49 dari skala 4,00, yang menunjukkan bahwa sistem dinilai mudah digunakan serta bermanfaat dalam membantu pengguna menemukan rekomendasi film yang sesuai dengan preferensinya.

Kata Kunci: Sistem Rekomendasi Film, Content-Based Filtering, Natural Language Processing, N-Gram, TF-IDF.

ABSTRACT

The rapid growth of streaming services has significantly increased the number of available movies, making it difficult for users to find movies that match their preferences. This study aims to develop a web-based movie recommendation system using the Content-Based Filtering (CBF) method by applying Natural Language Processing (NLP) techniques for text processing. The system was developed using the IMDb Movies Dataset obtained from Kaggle by utilizing movie overviews (synopses), genres, cast (*crew*), and movie titles as the primary features for similarity calculation, while the release year was used as a filtering criterion to refine recommendation results according to user preferences. The text processing stages consisted of text preprocessing, N-Gram representation (unigram and bigram), Term Frequency–Inverse Document Frequency (TF-IDF) weighting, Cosine Similarity calculation, and the combination of similarity scores using the average similarity method. The system was implemented using Python as the recommendation engine and the Laravel framework as the user interface. System performance was evaluated using the Precision@10 metric across several testing scenarios, resulting in an average score of 0.78, indicating that the system was able to provide movie recommendations relevant to user preferences. In addition, user evaluation was conducted using the Technology Acceptance Model (TAM), obtaining an overall average score of 3.49 out of 4.00, indicating that the proposed system was perceived as easy to use and useful in assisting users in finding movie recommendations that matched their preferences.

Keywords: Movie Recommendation System, Content-Based Filtering, Natural Language Processing, N-Gram, TF-IDF.