

ABSTRAK

Peningkatan jumlah publikasi ilmiah di Indonesia menimbulkan tantangan baru bagi mahasiswa dan peneliti dalam menemukan artikel yang relevan. Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan sistem rekomendasi artikel ilmiah berbasis *hybrid filtering*, yaitu penggabungan metode *content-based filtering* dan *collaborative filtering*. Sistem ini dirancang untuk membantu pengguna menemukan referensi yang sesuai dengan preferensi dan kebutuhan mereka, serta mengatasi masalah *cold start* yang umum terjadi pada sistem rekomendasi konvensional.

Metode yang digunakan dalam penelitian ini mencakup tahapan *preprocessing* data seperti *cleansing*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Pendekatan *content-based filtering* dilakukan dengan teknik TF-IDF dan *cosine similarity* pada judul dan abstrak artikel, sedangkan *collaborative filtering* menggunakan algoritma K-Nearest Neighbors (KNN) berbasis data rating pengguna. Dataset yang digunakan terdiri dari 1.012 data artikel dan 1.558 data interaksi pengguna. Evaluasi dilakukan menggunakan metrik *precision@k*, *recall@k*, dan *Mean Absolute Error* (MAE) untuk mengukur performa sistem.

Hasil pengujian menunjukkan bahwa model *hybrid filtering* memberikan performa terbaik dengan rata-rata *precision* sebesar 86,76% dan *recall* sebesar 80%, serta nilai MAE terendah pada $k = 23$. Hal ini menunjukkan bahwa penggabungan dua pendekatan mampu meningkatkan akurasi dan relevansi rekomendasi artikel ilmiah. Sistem ini diharapkan dapat menjadi solusi efektif dalam membantu proses pencarian referensi ilmiah yang relevan dan personal di lingkungan akademik.

Kata kunci: sistem rekomendasi, *hybrid filtering*, *content-based filtering*, *collaborative filtering*, artikel ilmiah.

ABSTRACT

The increasing number of scientific publications in Indonesia has created new challenges for students and researchers in finding relevant articles. To address this issue, this study developed a scientific article recommendation system using a hybrid filtering approach, which combines content-based filtering and collaborative filtering methods. The system is designed to help users find references that match their preferences and research needs while also addressing the cold start problem commonly found in conventional recommendation systems.

The method used in this research includes data preprocessing stages such as cleansing, case folding, tokenizing, stopword removal, and stemming. The content-based filtering approach uses TF-IDF and cosine similarity applied to article titles and abstracts, while collaborative filtering utilizes the K-Nearest Neighbors (KNN) algorithm based on user rating data. The dataset consists of 1,012 scientific articles and 1,558 user interaction records. Model evaluation was conducted using the metrics precision@k, recall@k, and Mean Absolute Error (MAE) to assess system performance.

The testing results show that the hybrid filtering model achieved the best performance, with an average precision of 86.76% and recall of 80%, and the lowest MAE value at $k = 23$. These results indicate that combining the two approaches improves the accuracy and relevance of the article recommendations. The proposed system is expected to serve as an effective solution in supporting personalized and relevant scientific reference searches in academic environments.

Keywords: *recommendation system, hybrid filtering, content-based filtering, collaborative filtering, scientific article.*