

ABSTRAK

Meningkatnya jumlah buku yang diterbitkan setiap tahun (mencapai lebih dari 2 juta judul secara global) semakin mempersulit pembaca dalam menemukan karya yang relevan dengan minat dan kebutuhan mereka, khususnya untuk buku berbahasa Indonesia. Sistem rekomendasi buku menjadi solusi dengan menyaring dan menampilkan daftar rekomendasi berdasarkan preferensi pengguna. Metode *Content-Based Filtering* (CBF) mampu merekomendasikan buku tanpa memerlukan parameter rating, namun sering menghasilkan daftar yang kurang beragam akibat *overspecialization*. Sebaliknya, *Collaborative Filtering* (CF) dapat meningkatkan keberagaman rekomendasi, tetapi rentan terhadap masalah *cold-start* jika data rating pengguna belum mencukupi. Oleh karena itu integrasi kedua pendekatan tersebut dalam kerangka *Hybrid Filtering* dipandang efektif untuk menyeimbangkan akurasi dan variasi rekomendasi. Namun *Hybrid Filtering* sendiri masih belum fleksibel dalam mengakomodasi kriteria multi-aspek seperti popularitas atau tahun terbit.

Penelitian ini mengusulkan sistem rekomendasi buku yang menggabungkan *Hybrid Filtering* (CBF dan CF) dengan metode multi-kriteria WASPAS (*Weighted Aggregated Sum Product Assessment*). Data sekunder yang digunakan diperoleh melalui *web scraping* Goodreads Indonesia Group sebanyak 2.820.808 data, terdiri dari 2 jenis data yaitu data buku dan data rating. *Preprocessing* data buku meliputi modifikasi tabel, *case folding*, *cleansing*, *tokenization*, *stopwords removal*, dan *stemming*. CBF dibangun dengan representasi TF-IDF dan *cosine similarity* untuk menghasilkan daftar kandidat awal berdasarkan kemiripan konten. CF diterapkan pada data rating dengan pembagian 80% data latih dan 20% data uji, menggunakan algoritma *cosine similarity* dan KNN untuk memprediksi skor rating *item*. Hasil rekomendasi CBF dan CF digabung secara linear untuk membentuk daftar kandidat awal. Selanjutnya pengguna memilih *genre* dan menetapkan bobot skala *Likert* pada kriteria rating, popularitas, kesamaan *genre*, dan tahun terbit. Bobot-bobot ini dinormalisasi, lalu diaplikasikan dalam WSM dan WPM pada metode WASPAS untuk menghitung skor akhir tiap buku. Sepuluh buku dengan skor tertinggi disajikan sebagai rekomendasi akhir.

Evaluasi model dilakukan dengan menggunakan *Intra-List Similarity* (ILS) untuk menilai keberagaman rekomendasi. Hasil penelitian menunjukkan bahwa nilai ILS terendah untuk top-10 dihasilkan oleh model *Hybrid* dengan WASPAS sebesar 0.326 yang menandakan bahwa hasil rekomendasinya lebih beragam dibandingkan ILS 0,590 pada top-10 rekomendasi *Hybrid* (CBF dan CF) saja. Dapat disimpulkan bahwa integrasi WASPAS berhasil meningkatkan variasi rekomendasi sehingga sistem yang dikembangkan mampu memberikan rekomendasi buku yang lebih relevan dan beragam bagi pembaca berbahasa Indonesia.

Kata Kunci: Sistem Rekomendasi, *Hybrid Filtering*, WASPAS

ABSTRACT

The increasing number of books published each year (reaching more than 2 million titles globally) makes it increasingly difficult for readers to find works that align with their interests and needs, especially for Indonesian-language books. A book recommendation system offers a solution by filtering and presenting a list of recommendations based on user preferences. Content-Based Filtering (CBF) can recommend books without requiring rating parameters, but it often produces lists that lack diversity due to overspecialization. In contrast, Collaborative Filtering (CF) can enhance recommendation diversity but is prone to cold-start issues when user rating data are insufficient. Therefore integrating these two approaches within a Hybrid Filtering framework is considered effective for balancing accuracy and diversity in recommendations. However Hybrid Filtering alone remains inflexible in accommodating multi-aspect criteria such as popularity or publication year.

This study proposes a book recommendation system that combines Hybrid Filtering (CBF and CF) with the multi-criteria WASPAS method (Weighted Aggregated Sum Product Assessment). Secondary data were obtained via web scraping from the Goodreads Indonesia Group, totaling 2,820,808 data records, it consists of two types of data: book data and rating data. Book data preprocessing included table modification, case folding, cleansing, tokenization, stopwords removal, and stemming. The CBF component was constructed using TF-IDF representation and cosine similarity to generate an initial candidate list based on content similarity. The CF component was applied to rating data split 80% for training and 20% for testing, using cosine similarity and K-Nearest Neighbors to predict item rating scores. The recommendation outputs from CBF and CF were linearly combined to form the initial candidate list. Next, users selected genres and assigned Likert-scale weights to the criteria of rating, popularity, genre similarity, and publication year. These weights were normalized and then applied within the WSM and WPM calculations of the WASPAS method to compute each book's final score. The ten books with the highest scores were presented as the final recommendations.

The evaluation model was conducted using Intra-List Similarity (ILS) to assess the diversity of recommendations. The results showed that the lowest ILS value for the top 10 generated by the Hybrid model with WASPAS was 0.326, indicating that the recommendation results were more diverse compared to the ILS of 0.590 for the top 10 recommendations of Hybrid (CBF and CF) alone. It can be concluded that the WASPAS integration successfully increased the variety of recommendations so that the developed system was able to provide more relevant and diverse book recommendations for Indonesian-speaking readers.

Keywords: Recommendation System, Hybrid Filtering, WASPAS