

## ABSTRACT

Research in the field of recommendation systems is currently growing rapidly. Recommendation system is a technology that is important in helping users to get a personalized recommendation for certain items. One approach that is often used in recommendation systems is collaborative filtering. Collaborative filtering is an algorithm that is base on prediction and recommendation of the assessment or the behavior of other users in the system. One of the challenges faced in collaborative filtering is a cold start. Cold start is a condition in which the recommendation system can not infer optimally because of new users and / or new items coming into the system. This causes information in the dataset is incomplete or sparse. Cold start could occur if there is a user and / or a new item that comes into the recommendation systems. There have been many studies conducted to deal with cold start. One of the popular ways is using ranking-oriented collaborative filtering approach like Borda-count and Copeland-score. But there is no studies have compared the performance and accuracy of the recommendation results between Borda-count and Copeland-score as ranking-oriented collaborative filtering approach. Evaluation is conducted by comparing the results of the performance and ranking accuracy of Borda-count and Copeland-score utilizing normalized discounted cumulative gain (NDCG) metric. From the results of the evaluation, the conclusion obtained that the Copeland-score method resulted in higher ranking accuracy of 1% compared to Borda-count. But based on the results of the significance test showed no significant difference. The Borda-count method runs about 28.000 times faster than the Copeland-score method. In addition, the clustering stage also affects the accuracy of the ranking predictions. In assisting with this research, Python library was built, especially for Borda-count and Copeland-score method.

**Keywords** – Collaborative Filtering, New User, Cold Start, Borda-count, Copeland-score and Significance.

## INTISARI

Penelitian di bidang sistem rekomendasi saat ini sedang berkembang pesat. Sistem rekomendasi adalah teknologi yang penting dalam membantu pengguna untuk mendapatkan rekomendasi pribadi untuk item tertentu. Salah satu pendekatan yang sering digunakan dalam sistem rekomendasi adalah *collaborative filtering*. *Collaborative filtering* adalah sebuah algoritme yang mendasarkan pada prediksi dan rekomendasi dari penilaian atau perilaku pengguna lain dalam sistem. Salah satu tantangan yang dihadapi dalam *collaborative filtering* adalah *cold start*. *Cold start* adalah suatu kondisi dimana sistem rekomendasi tidak dapat menyimpulkan secara optimal karena adanya pengguna baru dan/atau item baru yang masuk ke dalam sistem. Hal tersebut menyebabkan informasi dalam dataset tidak lengkap atau jarang. *Cold start* terjadi jika ada pengguna dan/atau item baru yang masuk ke dalam sistem rekomendasi. Ada banyak penelitian yang dilakukan untuk menangani masalah *cold start*. Salah satu cara populer adalah menggunakan pendekatan *collaborative filtering* berorientasi *ranking* seperti Borda-count dan Copeland-score. Tetapi penelitian saat ini belum ada yang membandingkan kinerja dan hasil rekomendasi pendekatan Borda-count dan Copeland-score, sebagai metode *collaborative filtering* berorientasi *ranking*. Tahapan evaluasi dilakukan dengan membandingkan hasil dari kinerja dan hasil rekomendasi dari Borda-count dan Copeland-score memanfaatkan *normalized discounted cumulative gain* (NDCG) matriks. Dari hasil evaluasi, kesimpulan yang didapat adalah bahwa algoritme Copeland-score menghasilkan *ranking* lebih tinggi 1% dibandingkan Borda-count. Tetapi berdasarkan hasil uji signifikansi menunjukan perbedaan yang tidak signifikan. Algoritme Borda-count berjalan sekitar 28.000 kali lebih cepat dibandingkan dengan algoritme Copeland-score. Selain itu tahapan *clustering* yang dilakukan juga berpengaruh terhadap nilai prediksi *ranking*. Dalam membantu pelaksanaan penelitian ini dikembangkan modul Python, khususnya untuk algoritme Borda-count dan Copeland-score.

**Kata kunci** – *Collaborative Filtering*, Pengguna Baru, *Cold Start*, Borda-count, Copeland-score dan Signifikansi.