

Perancangan dan analisis sistem rekomendasi berdasarkan rating movielens 100k dengan menggunakan metode collaborative filtering dan mapreduce = Implementation and analysis of recommender system based on rating movielens 100k using collaborative filtering methods and mapreduce

Billy Surya Putra, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20457086&lokasi=lokal>

Abstrak

Sistem rekomendasi adalah sebuah teknik untuk menyediakan saran terkait suatu hal yang dapat dimanfaatkan oleh pengguna. Saran dapat berupa produk maupun jasa yang ditawarkan. Saran yang diberikan adalah produk atau jasa yang belum pernah digunakan atau dibeli oleh pengguna tersebut. Sistem rekomendasi, khususnya dengan menggunakan K-Nearest Neighbor KNN, mencapai kesuksesan pada beberapa akhir tahun ini.

Penelitian ini akan diimplementasikan K-Nearest Neighbor pada komputasi terdistribusi yaitu MapReduce untuk merancang sistem rekomendasi dengan menggunakan Item Based Collaborative Filtering IBCF dan User Based Collaborative Filtering UBCF pada dataset Movielens 100k. Penelitian akan menggunakan beberapa komputasi penghitung kesamaan yaitu Cosine Based Similarity, Pearson Correlation Similarity dan Euclidean Distance.

Hasil percobaan yang didapat adalah algoritma Euclidean Distance menghasilkan performa terbaik dalam waktu proses dan nilai keakuratan. Pada pendekatan IBCF, Euclidean Distance membutuhkan waktu proses dengan rata-rata 13 sekon dan nilai korelasi sebesar 0.84. Sedangkan pada UBCF, Euclidean Distance membutuhkan waktu proses dengan rata-rata 32 sekon dan nilai korelasi sebesar 0.84.

.....Recommender system is a technique to provide suggestions related to a thing that can be used for user. Suggestions can be products and services offered. The advice given is a product or service that has never been used or purchase by the user. The recommendation system, especially by using K Nearest Neighbor KNN, achieving success in several year.

This research will be implemented K Nearest Neighbor at distributed process that called MapReduce to arrange system by using Item Based Collaborative Filtering IBCF and User Based Collaborative Filtering UBCF on Movielens 100k dataset. The research will use several techniques to compute similarities such as Cosine Based Similarity, Pearson Correlation Similarity and Euclidean Distance.

The result of the experiment is Euclidean Distance algorithm give the best performance in process time and correlation. In the IBCF approach, Euclidean Distance takes process around 13 seconds and correlation value is 0.84. And at UBCF, Euclidean Distance takes processing time around 32 seconds and correlation value is 0.84.