

Abstrak

Recommender system berbasis *collaborative filtering* merupakan suatu sistem yang dapat memberikan rekomendasi berupa prediksi rating terhadap suatu *item* berdasarkan persamaan karakteristik atau kecenderungan *user* dalam merating suatu *item*. Sistem ini telah banyak digunakan baik oleh perusahaan atau non-profit organisasi. Sehingga, diperlukan *recommender system* yang dapat menangani data dalam skala besar.

ClustKNN merupakan algoritma yang sederhana dan intuitive yang cukup cocok untuk menangani dataset yang besar, dengan waktu prediksi yang relative lebih cepat dari pada *user-based KNN* biasa. Tugas akhir ini menganalisis akurasi prediksi *rating* dan waktu prediksi yang dihasilkan oleh *recommender system* setelah diimplementasikan algoritma *ClustKNN* dan *user-based KNN*. Parameter yang digunakan dalam analisis adalah perbandingan *training set* dengan *test set*, jumlah *cluster*, serta ukuran *neighbor*.

Pada algoritma *ClustKNN*, kecenderungan *active user* dan kecenderungan *user* lain dalam merating suatu *item* sangat menentukan hasil prediksi rating. Akurasi prediksi yang dihasilkan oleh algoritma *ClustKNN* dan *user-based KNN* akan semakin meningkat dengan bertambahnya jumlah data yang terdapat pada *training set*, selain itu banyaknya jumlah *cluster* dan ukuran *neighbor* juga sangat berpengaruh, semakin banyak jumlah *cluster* dan besar ukuran *neighbor* maka akurasi akan semakin baik.

Kata kunci: *recommender system*, *user-based collaborative filtering*, algoritma *ClustKNN*, algoritma *user-based KNN*.