METODE RECTREE PADA RECOMMENDER SYSTEM BERBASIS COLLABORATIVE FILTERING

Ali Murtado Fauzarrohman[1], Yanuar Firdaus A.w.[2], Warih Maharani[3]

[1]Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Recommender system adalah sebuah aplikasi yang dapat memberikan rekomendasi berupa prediksi rating terhadap suatu item berdasarkan persamaan karakteristik user dalam memberikan informasi.

Tugas akhir ini mengimplementasikan dan menganalisis user-based collaborative filtering pada recommender system, yang menerapkan algoritma CorrCF dan metode RecTree. Tugas akhir ini menganalisis kecepatan untuk menentukan prediksi dan akurasi prediksi rating yang dihasilkan oleh recommender system setelah diimplementasikan algoritma CorrCF dan metode RecTree. Parameter yang digunakan dalam analisis adalah parameter untuk metode RecTree serta perbandingan training set dengan test set.

Pada metode RecTree, active user history serta keberadaan user lain yang mirip dan pernah merating item sangat menentukan hasil dari prediksi rating untuk active user. Akurasi prediksi yang dihasilkan oleh algoritma CorrCF dan metode rectree akan semakin meningkat dengan bertambahnya jumlah data yang terdapat pada training set.

Waktu yang dibutuhkan oleh metode RecTree jauh lebih cepat dibandingkan algoritma CorrCF dalam menentukan prediksi. Selain itu, dengan pemilihan parameter yang tepat maka akurasi yang dihasilkan oleh metode RecTree lebih baik dibandingkan algoritma CorrCF yang ditunjukkan oleh nilai MAE yang dihasilkan oleh metode RecTree lebih kecil daripada algoritma CorrCF.

Kata Kunci : recommender system, collaborative filtering, algoritma CorrCF, metode RecTree

Abstract

Recommender system is an application that can give a recommendation in term of rating prediction of an item, based on the similarity of user characteristic in giving information.

In this final project, the implementation and the analysis of the user-based collaborative filtering recommender system, which applies CorrCF algorithm and RecTree method, is performed. Initially, CorrCF algorithm and RecTree method is implemented into the recommender system. Then, the analysis is carried out to the speed and the accuracy of rating prediction result that is given by the recommender system. Comparison parameter on training set and set test is used in the analysis.

In the RecTree method, active user history and the existence of other similar users which have rated items, determine the result of a predicted rating for the active user. The accuracy of prediction, which is resulted by both algorithm, increases with the increase of the number of data in the training set.

Time needed by the RecTree method is much faster than CorrCF algorithm to determine prediction. In addition, with appropriate selection of parameter then the accuracy produced by the RecTree method is better than the CorrCF algorithm which is shown by the MAE value generated by RecTree method is smaller than generated by CorrCF algorithm.

Keywords : recommender system, collaborative filtering, CorrCF algorithm, RecTree method

BAB I Pendahuluan

1.1 Latar belakang masalah

Saat ini banyak sekali data yang beredar di internet[10]. Data-data tersebut sangat bermacam-macam, seperti teks, video, gambar, dan lain-lain. Hal ini menyebabkan dibutuhkannya suatu alat bantu pencarian yang handal untuk mendapatkan data atau informasi yang diinginkan.

Recommender System adalah sebuah sistem pendukung penyaringan informasi (*Information Filtering*) yang menampilkan rekomendasi sebuah produk (baik musik, buku, film, gambar, dsb) yang mungkin menarik bagi seorang *user*[12]. Secara umum *Recommender system* dibagi menjadi empat, yaitu *Content-Based Filtering System*, *Recommendation Support System*, *Social Data Mining System*, dan *Collaborative Filtering System*[11]. Teknik yang paling umum digunakan dalam memberikan informasi dimana melibatkan kolaborasi antara beberapa pengguna, beberapa sudut pandang, dan source serta kondisi dimana tersedia informasi *history* dari pengguna adalah teknik *Collaborative*[11]. Teknik ini memberikan rekomendasi kepada pengguna berdasarkan kemiripan selera dengan pengguna lain. Teknik ini biasa ditemui pada situs-situs yang menyediakan layanan rekomendasi musik , video, dan film, seperti pada movielens.org. Oleh karena itu, *collaborative filtering* digunakan pada penyusunan tugas akhir ini untuk melakukan proses filtering terhadap *dataset* yang digunakan.

Salah satu metode berbasis *collaborative filtering* yang telah terbukti tingkat keakuratannya adalah *correlation collaborative filterin*g atau lebih dikenal dengan sebutan *CorrCF*[6]. *CorrCF* menggunakan *pearson's correlation* untuk menentukan *similarity* dan *deviation from mean* untuk menentukan prediksi. Namun *CorrCF* memiliki kelemahan yaitu waktu yang lama untuk menentukan prediksi.

Metode yang lain yang lebih cepat dalam menentukan prediksi dan memiliki tingkat keakuratan yang baik adalah *RecTree*[3][7]. *RecTree* sendiri merupakan akronim dari struktur data dan algoritma untuk *collaborative filtering* yang disebut dengan *RECommendation Tree* [6][7]. Metode *RecTree* menggunakan suatu sistem rangka kerja yang terdiri atas tiga fase, yaitu: mempartisi dataset dalam kelompok-kelompok, melatih masing-masing kelompok dengan *collaborative filtering* untuk mendapat bobot kemiripan antar *user*, dan melakukan perhitungan prediksi[6]. Metode *RecTree* membagi dataset menjadi dua secara rekursi dengan *KMeans+*.

Dalam penyusunan tugas akhir ini dilakukan implementasi algoritma *CorrCF* dan metode *RecTree* pada *recommender system* berbasis *collaborative filtering*. Analisis yang dilakukan adalah mengenai perbandingan antara algoritma *CorrCF* dan metode *RecTree* dilihat dari sisi kecepatan waktu penentuan prediksi dan

keakuratan hasil prediksi. Dalam hal ini algoritma *CorrCF* hanya digunakan sebagai pembanding terhadap kinerja dari metode *RecTree*.

1.2 Perumusan masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka permasalahan yang dijabarkan dan diteliti antara lain:

1. Bagaimana menerapkan algoritma *CorrCF* dan metode *RecTree* pada *Recommender System* berbasis *collaborative filtering*.
2. Bagaimana perbandingan antara algoritma *CorrCF* dan metode *RecTree* dilihat dari kecepatan waktu penentuan prediksi.
3. Bagaimana perbandingan antara algoritma *CorrCF* dan metode *RecTree* dilihat dari ketepatan prediksi yang dihasilkan berdasarkan banyaknya training set dan tes set.

Adapun batasan masalah tugas akhir ini adalah sebagai berikut :

1. Metode yang digunakan dalam penentuan rekomendasi adalah algortima *CorrCF* dan metode *RecTree*.
2. Dataset yang direkomendasikan adalah dataset film yang berasal dari MovieLens.org.

1.3 Tujuan

Tujuan yang ingin dicapai dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

1. Mengimplementasikan algoritma *CorrCF* dan metode *RecTree* pada *Recommender System* berbasis *collaborative filtering*
2. Mengetahui perbandingan antara algoritma *CorrCF* dan metode *RecTree* dilihat dari kecepatan waktu eksekusi untuk mendapatkan hasil prediksi yang diinginkan.
3. Mengetahui perbandingan antara algoritma *CorrCF* dan metode *RecTree* dilihat dari ketepatan prediksi yang dihasilkan berdasarkan banyaknya *training set* dan *tes set*.

1.4 Metodologi penyelesaian masalah

Metodologi yang digunakan dalam memecahkan masalah di atas adalah dengan menggunakan langkah-langkah berikut:

1. Studi Literatur
   Pencarian referensi dan sumber-sumber yang berhubungan dengan *Recommender System*, *Collaborative Filtering,* algoritma *CorrCF* dan metode *RecTree* dalam menyelesaikan tugas akhir ini.
2. Pengumpulan data
   Mengumpulkan data produk yang diperlukan sebagai *training set* dan *test set.*
3. Analisis dan Perancangan Perangkat Lunak

Melakukan analisis dan perancangan terhadap perangkat yang dibangun, menganalisis metode yang digunakan untuk menyelesaikan permasalahan, termasuk menentukan bahasa pemrograman yang digunakan, arsitektur, fungsionalitas, dan antarmuka sistem.

4. Implementasi dan Pembangunan Perangkat Lunak
   Implementasi diawali dengan pembuatan basis data untuk menyimpan data yang berasal dari movielens.org. Lalu dilakukan pembuatan perangkat lunak dengan menerapkan algoritma *CorrCF* dan metode *RecTree*. Kemudian membuat interface aplikasi untuk memudahkan user dalam mengakses perangkat lunak.
5. Testing & Analisis
   Pengujian perangkat lunak dibagi menjadi dua macam, yaitu :
   a. Pengujian untuk mengetahui kecepatan waktu eksekusi untuk mendapatkan hasil prediksi yang dinginkan.
   b. Pengujian untuk mengetahui ketepatan hasil rekomendasi yang dihasilkan berdasarkan banyaknya *training set* dan *tes set* melalui MAE (*Mean Absolute Error*).

   Setelah itu maka dilakukan analisis terhadap kecepatan dan ketepatan dari algoritma *CorrCF* dan metode *RecTree* berdasarkan hasil pengujian.
6. Pengambilan Kesimpulan dan penyusunan laporan Tugas Akhir.

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil uji coba dan analisis yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Algoritma *CorrCF* dan metode *RecTree* dapat diimplementasikan pada *recommender system* berbasis *collaborative filtering*.
2. Waktu yang dibutuhkan untuk menentukan prediksi dengan menggunakan metode *RecTree* jauh lebih cepat (kurang dari sepertiga) dibanding algoritma *CorrCF* karena pada metode *RecTree* data telah dibagi-bagi menjadi kelompok yang lebih kecil.
3. Dengan pemilihan nilai parameter β yang tepat (20-25% dari jumlah test set) maka ketepatan prediksi *rating* pada metode *RecTree* lebih baik dibandingkan algoritma *CorrCF* yang ditunjukkan oleh nilai MAE yang lebih kecil. Hal ini disebabkan pada metode RecTree dilakukan perbaikan pada penghitungan bobot *similarity*.

5.2 Saran

Setelah tugas akhir ini selesai dilakukan dan telah dilakukan analisis, penulis memiliki beberapa saran sebagai berikut:

1. Untuk menghasilkan rekomendasi yang sesuai dengan dengan tipe atau konten dari item yang direkomendasikan, sebaiknya menggunakan pendekatan *content-based filtering*.
2. Karena CoFE sudah tidak dikembangkan lebih lanjut, maka sebaiknya digunakan tools lain yang menggunakan metode *collaborative filtering*, seperti *Vogoo* dan *Taste*.

Daftar Pustaka

[1] Billsus, D. Pazzani, Michael J. 1998. *Learning Collaborative Information Filters*. Diunduh pada: http://www.ics.uci.edu/~pazzani/Publications/MLC98.pdf, 20 Januari 2009

[2] CoFE, http://eecs.oregonstate.edu/iis/CoFE/

[3] Kolleher, Jerome & Bridge, Derek. *RecTree Centroid: An Accurate, Scalable Collaborative Recommender*. Diunduh pada: http://www.cs.ucc.ie/~dgb/papers/Kelleher-Bridge-2003.pdf, 20 Januari 2009.

[4] LI, Ying-Ji, dkk. *A Hybrid Order Markov Tree Recommendation Algorithm*. Diunduh pada: http://68.153.197.235/CSREA/ICC2069.pdf, 20 Januari 2009.

[5] McRae, J., Piatek, A., Langley, A., Collaborative Filtering. 2004. Diunduh pada: http://www.imperialviolet.org/suprema.pdf, 20 Januari 2009.

[6] S. H. S. Chee. *Rectree:A linear collaborative filtering algorithm*. Master's thesis, Simon Fraser University, November 2000. Diunduh pada: http://fas.sfu.ca/pub/cs/theses/2000/SonnyHanMengCheeMSc.pdf , 20 Januari 2009.

[7] S. H. S.g Chee, J. H., and K. Wang. *Rectree:An efficient collaborative filtering method*. Lecture Notes in Computer Science, 2114, 2001.Diunduh pada: http://www.cs.sfu.ca/~wangk/pub/dawakfin.pdf, 20 Januari 2009

[8] Sarwar, B., Karypis,G., Konstan, J., Riedl, J. *Item-based Collaborative Filtering Recommendation Algorithm*. 2001. Diunduh pada: http://www.grouplens.org/papers/pdf/www10_sarwar.pdf, 20 Januari 2009.

[9] Terveen, Loren & Will Hill. *Beyond Recommender Systems: Helping People Help Each Other*. Dalam *HCI In The New Millennium*, Jack Carroll, ed., Addison-Wesley, 2001. Diunduh pada: http://www.grouplens.org/papers/pdf/rec-sys-overview.pdf , 20 Januari 2009.

[10] Vozalis, E. Margaritis, K. G. 2003. *Analysis of Recommender Systems Algorithms*.

[11] Wikipedia, *Collaborative Filtering*, 2003. Diunduh pada: http://en.wikipedia.com/wiki/collaborative_filtering, 20 Januari 2009.

[12] Wikipedia. *Recommender System*. Diunduh pada: http://en.wikipedia.org/wiki/Recommender_System , 20 Januari 2009

Telkom University