

IMPLEMENTASI DAN ANALISIS NAIVE BAYES CLASSIFIER DAN ITEM BASED COLLABORATIVE FILTERING PADA SWITCHING HYBRID RECOMMENDER SYSTEM

Septiwi Fitriana Dewi¹, Ema Rachmawati², Arfi Fitrandi³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Recommender system merupakan aplikasi yang memberikan prediksi terhadap suatu item kepada user berdasarkan karakteristik dari user dalam memberi informasi.

Tugas akhir ini mengimplementasikan dan menganalisis metode Switching Hybrid pada Recommender System dengan Naive Bayes Classifier dan Item Based Collaborative Filtering. Tugas akhir ini menganalisis pengaruh yaitu sparsity problem terhadap akurasi prediksi rating yang dihasilkan oleh recommender system setelah diimplementasikan metode Switching Hybrid dengan Naive Bayes Classifier dan Item Based Collaborative Filtering.

Penggunaan metode Switching Hybrid pada recommender system dapat membantu dalam meningkatkan tingkat akurasi dengan memanfaatkan kelebihan dari kedua metode yang digunakan yaitu naive bayes classifier dan Item Based Collaborative Filtering. Kedua algoritma ini memiliki nilai komputasi yang tinggi dalam penghitungan prediksi oleh karena itu akan dihasilkan tingkat akurasi yang tinggi dan akan memberikan suatu hasil prediksi jika salah satu metode gagal dalam melakukan pembuatan hasil prediksi. Hal ini disebabkan karena kelemahan dari metode yang digunakan yaitu data yang sparse akan membuat hasil prediksi dari Item Based CF menjadi 0 dan tidak akurat oleh karena itu digabungkan dengan Naive Bayes Classifier yang dapat menangani sparsity problem sehingga akan ada nilai prediksi. Untuk Pemilihan prediksi menggunakan nilai α dan β , α adalah selisih antar nilai probabilitas dari Naive Bayes Classifier yang telah diurutkan dan β adalah selisih nilai prediksi dari Item Based Collaborative Filtering dengan Naive Bayes Classifier.

Hasil pengujian menunjukkan bahwa semakin banyak jumlah neighborhood, maka akurasi prediksi yang dihasilkan semakin baik. Performansi terbaik terjadi pada saat nilai α dan β yang digunakan adalah nilai optimal ($\alpha=0.35$ dan $\beta=0.8$).

Kata Kunci : recommender system, switching hybrid, item based collaborative filtering, naive bayes classifier, sparsity

Telkom
University

Abstract

Recommender system is an application giving prediction of an item based to user based on user's characteristics in giving information.

This final assignment implements and analyzes Naive Bayes Classifier and Item Based Collaborative Filtering in Switching Hybrid Recommender System.

This final assignment is to analyze influence of sparsity to the accuracy of the rating prediction generated by recommender system after the implementation of Switching Hybrid method with Naive Bayes Classifier and Item Based Collaborative Filtering.

Switching Hybrid method in Recommender System can assist improving accuracy by exploiting the advantages of both method , Naive Bayes Classifier and Item Based Collaborative Filtering. Both of these algorithm have a high performance computing value prediction calculation so can produce high degree accuracy and the algorithms will provide an outcome prediction if one method fails in making the manufacture of predicted results. It's Happen because the weakness of the method used here is sparse data that would make the predistid results of item based CF is zero and inaccurate so combined with naive bayes classifier is handle sparsity problem and give value prediction. For selection of prediction using value α and value β , α is the difference between the value of the probability of the Naïve Bayes that have been sequenced and the difference between the predicted value of Item Based Collaborative Filtering with Naïve Bayes Classifier.

Test Results showed that more the number of neighborhood will give resulting good prediction accuracy. The best performance occurs when value of α and β used is the optimal value($\alpha=0.35$ and $\beta=0.8$).

Keywords : Switching Hybrid, Item Based Collaborative Filtering, Naive Bayes Classifier, Scalability, Sparsity

1 Pendahuluan

1.1 Latar belakang

Pada saat sekarang ini terdapat banyak peningkatan dalam penyediaan informasi, ada yang bersifat digital, elektronik maupun *online*. Dengan adanya informasi dengan jumlah yang berlebihan menimbulkan masalah yang sangat potensial, yaitu seperti permasalahan dalam penyaringan dan pengefisienan pemberian informasi yang relevan kepada pengguna. Serta adanya permasalahan *information overload* yaitu banyaknya informasi yang benar yang muncul padahal lebih baik informasi yang muncul itu yang lebih diprioritaskan yang dicari oleh pengguna. *Internet Search Engines* pada umumnya berpotensi untuk menemukan informasi sesuai *query* inputan dari pengguna, tetapi kemungkinan yang sesuai yang pengguna inginkan itu sangat kecil sekali. Oleh karena itu diperlukan sebuah sistem yang dapat menyortir tentang sebuah ekstraksi informasi dan juga dapat memprediksikan informasi yang sesuai dengan keinginan pengguna dari kata kunci yang diberikan. Sistem ini disebut *recommender system*, sistem ini dapat memberikan prediksi terhadap *item* baru dengan memperhatikan tingkat akurasi kesukaan pengguna terhadap *item* ini atau dapat juga berdasarkan referensi pengguna serta perilaku pengguna dan informasi [9].

Salah satu *recommender system* yaitu tentang *recommender system* buku yaitu sistem yang memberikan rekomendasi buku kepada *user* berdasar perilakunya. Dengan adanya *recommender system* buku, maka dapat dengan mudah menganalisis perilaku *user*, sehingga kita akan dengan mudah mendapatkan buku yang sesuai dengan *user*.

Di dalam pembuatan *recommender system* kita harus memperhatikan perilaku *user* yang mungkin mempunyai *similarity* kesamaan satu sama lain. Selain itu kita juga harus memperhatikan isi dari buku tersebut, biasanya seorang *user* akan suka dengan buku tersebut karena tertarik dengan isi dari buku yang dia baca. Sebuah buku biasanya dibaca oleh banyak *user*, dan ada *user* yang hanya sekedar membaca, tidak memberikan *rating* kepada semua *item* buku dari yang dia baca sehingga mengakibatkan adanya permasalahan *sparsity problem*[12]. Oleh karena itu kita perlu menggunakan teknik-teknik dalam membangun sebuah *recommender system* buku yang dapat memecahkan permasalahan tersebut.

Di dalam *pembangunan recommender system* ini digunakan teknik jenis *hybrid* yaitu *switching*, karena dengan jenis *hybrid* ini maka sistem akan dapat memilih salah satu teknik yang dihybridkan sesuai dengan kriteria yang ada, yaitu berdasar kriteria dengan mengukur tingkat *confidence* dari kedua teknik dalam menghasilkan prediksi dengan memanfaatkan kelemahan dan keunggulan dari teknik tersebut. Pada tugas akhir ini metode yang dikombinasikan dalam jenis *switching hybrid* adalah *item based collaborative filtering* dan *naive bayes classifier*. Algoritma tersebut dapat menangani data yang tidak lengkap sehingga dapat memberikan solusi

dalam pembangunan *recommender system* buku dengan tingkat akurasi yang bagus dari permasalahan yang ada.

1.2 Perumusan masalah

Berdasarkan latar belakang permasalahan, maka masalah yang akan dirumuskan adalah sebagai berikut:

- a. Bagaimana membangun *recommender system* dengan menggunakan *switching hybrid recommender system* berbasis *item based collaborative filtering* dan *naïve bayes classifier*?
- b. Bagaimana akurasi prediksi rating sistem rekomendasi *switching hybrid recommender system* berbasis *item based collaborative filtering* dan *naïve bayes classifier*?

Adapun batasan masalah yang ada dalam tugas akhir ini adalah sebagai berikut:

- a. *Item* yang dijadikan objek rekomendasi adalah buku yang telah dirating oleh user.
- b. *Dataset* yang digunakan berasal dari *BookCrossing*.
- c. *Feature item* yang digunakan adalah *genre*, *publisher*, *author*.
- d. Tidak membahas masalah tingkat kepuasan *user* terhadap hasil prediksi.

1.3 Tujuan

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

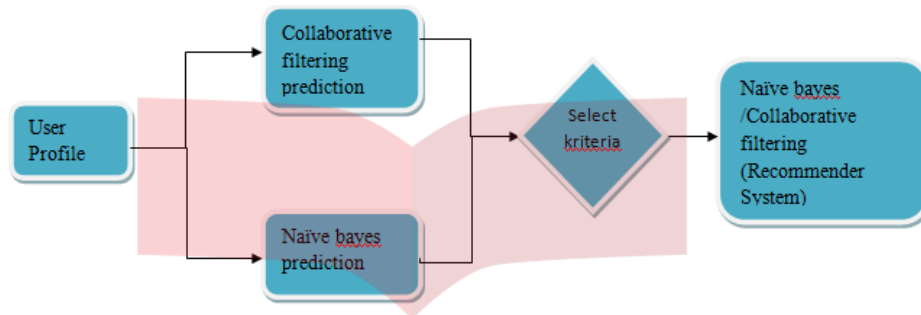
- a. Membangun teknik *switching hybrid recommender system* berbasis *item based collaborative filtering* dan *naïve bayes classifier* pada *recommender system*.
- b. Menganalisis akurasi hasil prediksi *recommender system* yang menerapkan *switching hybrid recommender system* berbasis *item based collaborative filtering* dan *naïve bayes classifier*.

1.4 Metodologi penyelesaian masalah

Metodologi yang digunakan dan langkah-langkah dalam penyelesaian masalah yang telah dirumuskan di atas adalah:

- a. Studi Literatur.
Merupakan tahapan dalam mempelajari konsep dan teori pendukung untuk memecahkan permasalahan. Dalam tugas akhir ini, studi *literatur meliputi pembelajaran konsep recommender system, collaborative filtering, naïve bayes classifier, switching hybrid* serta informasi lainnya yang menunjang pembuatan tugas akhir ini.
- b. Perancangan dan Pemodelan Sistem.

Pada tahap ini akan dirancang sebuah model *recommender system* yang dibangun dengan menggunakan metode *switching hybrid recommender system*. Metode ini dibangun dengan menggunakan dua pendekatan, yaitu *collaborative filtering* dan *naïve bayes classifier* yang akan digunakan untuk menentukan prediksi.



Gambar 1-1 Perancangan *switching hybrid*

- c. Implementasi Sistem.
Pada tahap ini model yang sudah dirancang kemudian diimplementasikan. Pengumpulan *dataset* dilakukan pada tahap ini, dataset yang dipakai berasal dari database *BookCrossing*. Implementasi ini akan menentukan prediksi yang dilakukan dengan *switching* antara *item based collaborative filtering* dan *naïve bayes classifier*.
- d. Pengujian Sistem dan Analisis.
Recommender system yang telah diimplementasikan dengan menggunakan *item based collaborative filtering* dan *naïve bayes classifier* dengan teknik *switching hybrid* diuji menggunakan dataset buku dari *BookCrossing*. Pengujian ini dilakukan dengan menghitung nilai *Mean Absolute Error* (MAE) yang dihasilkan oleh sistem yang akan dianalisis akurasi.
- e. Perumusan Kesimpulan dan Penyusunan Buku.
Di tahap ini akan dilakukan perumusan kesimpulan berdasarkan analisis yang telah dilakukan terhadap sistem. Kemudian akan dilakukan penyusunan buku tugas akhir dan pengumpulan dokumentasi dengan mengikuti kaidah yang telah ditetapkan oleh institusi.

Telkom
University

5 Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil pengujian dengan metode *Naive Bayes Classifier* dan *Item Based Collaborative Filtering* pada *Switching Hybrid Recommender System* maka dapat disimpulkan bahwa :

- a. Dengan pemilihan nilai parameter α (selisih antar nilai probabilitas kelas yang telah diurutkan dalam *naïve bayes classifier*) yang tepat ($\alpha = 0.35$) metode *switching hybrid* dengan *naïve bayes classifier* dan *item based collaborative filtering* dapat menghasilkan prediksi sesuai preferensi *user* dengan nilai MAE sekitar 0.700001.
- b. Dengan pemilihan nilai parameter β (selisih nilai prediksi dari *item based collaborative filtering* dan *naïve bayes classifier*) yang tepat ($\beta = 0.8$) metode *switching hybrid* dengan *naïve bayes classifier* dan *item based collaborative filtering* dapat menghasilkan prediksi sesuai preferensi *user* dengan nilai MAE sekitar 0.700001.
- c. Dengan pemilihan *feature* yang representatif akan mempengaruhi tingkat keakuratan hasil prediksi. *Feature genre* yang merupakan *feature* buku yang representatif dengan hasil prediksi sesuai preferensi *user* dengan MAE =0.725001.
- d. Semakin banyak jumlah *neighborhood*, maka akurasi prediksi yang dihasilkan oleh *recommender system* dengan metode *naive bayes classifier* dan *item based collaborative filtering* pada *switching hybrid* akan semakin tinggi. Hal ini disebabkan karena nilai *neighborhood* digunakan untuk menentukan jumlah *item* yang memiliki nilai *similarity* paling tinggi terhadap suatu *item* pada metode *item based collaborative filtering*.
- e. *Recommender system* dengan metode *naive bayes classifier* dan *item based collaborative filtering* pada *switching hybrid* dapat mengatasi kasus *sparsity problem*. Hal ini dikarenakan pada metode *naive bayes classifier* dan *item based collaborative filtering* pada *switching hybrid*, *item* yang belum dirating *user* lain, akan dihitung prediksi *ratingnya* menggunakan *naïve bayes classifier*.

5.2 Saran

- a. Untuk pengembangan selanjutnya, dalam pemilihan metode untuk klasifikasi penentuan prediksi bisa menggunakan selain *naïve bayes classifier* yang memiliki kemampuan lebih bagus seperti *Support Vector Machines*.
- b. Untuk dapat membuat akurasi prediksi yang lebih baik lagi dapat menggunakan pemilihan *feature* yang lebih representatif lagi seperti resensi buku.

Daftar Pustaka

- [1] Ahmad, Erival. *Klasifikasi Sms Pada Server Content Provider Menggunakan Metode Naïve Bayesian'*, Tugas Akhir, Institut Teknologi Telkom.2006.
- [2] Burke, Robin. *Hybrid Web Recommender System*. 2007.
- [3] Fisher, Danyel., Hildrum, Kris., Hong, Jason., Newman, Mark., Thomas, Megan., Vuduc, Rich. 2000. *SWAMI : A Framework for Collaborative Filtering Algorithm Evaluation and Development*.
- [4] Ghazanfar, Mutansar Ali, Adam Prugel-Bennett. "An Improved Switching Hybrid Recommender System Using Naive Bayes , *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. I, 2010.
- [5] Gunawardana, Asela, Guy Shani. 2009. *A survey of accuracy evaluation metrics of recommendation tasks. Journal of Machine Learning Research*, 10 (December 2009): 2935-2962.
- [6] Hapsari, Widi, Budi Susanto, Tantiny. *Klasifikasi Email dengan menggunakan metode Naïve Bayesian*. *Jurnal Informatika*, vol 3, 2007.
- [7] Mao, Ha, I. King, and M. Lyu, "Effective missing data prediction for collaborative filtering,". *Proceedings of the 30th annual international ACM SIGIR conferenceon Research and development in information retrieval*. ACM, 2007.
- [8] Melville, Prem, R. J. Mooney, and R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations, " *Proceedings of the Eighteenth National Conference on Artificial Intelligence(AAAI-2002)*,pp. 187-192, 2002.
- [9] Mobasher, Bamshad. *Recommender systems*. 2007.
- [10] Sarwar, Badrul, George Karypis, Joseph Konstan, John Riedl. *Item-Based Collaborative Filtering Recommendation Algorithms*. Minneapolis: University of Minnesota. 2001.
- [11] Shani, Guy, David Heckerman, Ronen I. Brafman. *An MDP-Based Recommender System. Journal of Machine Learning Research*, 2005: 1265~1295

- [12] Sieg, Ahu, Bamshad Mobasher, Robin Burke. *Improving the Effectiveness of Collaborative Recommendation with Ontology-Based User.2010 Profiles*. DePaul University. 2010.
- [13] Sinha, Rashmi. 2006. *Design Strategies for Recommender System*
- [14] Vozalis, Emmanouil, Konstantinos G. Margaritis. *Analysis of Recommender Systems Algorithms*. 2003.
- [15] W. Cohen, William. *Collaborative Filtering: A Tutorial*. 2006.
- [16] Zhang, H. and Su, Jiang. 2001. *Naïve Bayesian Classifiers for Ranking*.

