

ANALISIS DAN IMPLEMENTASI ALGORITMA BI-POLAR SLOPE ONE PADA PREDIKSI RATING BERBASIS COLLABORATIVE FILTERING

Aditya Pradana¹, Tjokorda Agung Budi Wirayuda², Ade Romadhony³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Dengan kemajuan teknologi informasi dan komunikasi, telah membuat perubahan di segala aspek kehidupan. Dengan memanfaatkan teknologi informasi dan komunikasi maka setiap orang dapat dengan mudah mendapatkan informasi. Namun hal ini akan menjadi masalah bila informasi yang kita dapat merupakan informasi yang tidak kita perlukan.

Tugas akhir ini bertujuan untuk menganalisis hasil prediksi rating yang dihasilkan oleh recommender system berbasis collaborative filtering yang menggunakan algoritma bi-polar slope one pada proses perhitungan rating. Tugas akhir ini menggunakan tiga teknik pembagian data pada implementasi algoritma. Yaitu membagi data dengan mean user, nilai yang ditentukan dan membiarkan sistem menentukan dimana data akan dibagi, teknik ini akan disebut auto balance.

Hasil prediksi yang dihasilkan oleh algoritma bi-polar slope one dengan teknik auto balance menunjukkan hasil yang lebih baik dibanding dua teknik lainnya pada data yang persebarannya merata.

Kata Kunci : recommender system, collaborative filtering, algoritma slope one

Abstract

The advance of communication and information technology have changed every aspect of live. By using information and communication technology everyone can have easy access to information. But it will be a problem if the information we got isn't the one we need.

The purpose of this project is to analyze the result of prediction made by recommender system based on collaborative filtering using bi-polar slope one algorithm to predict the rating. In this project we will use three data splitting technique. The first technique is to split the data using user's mean. The second one is to split the data using a value that was decided before. And the last one is to let the system split the data so that it will be balanced, this technique is called auto balance.

Prediction accuracy produced by auto balance shows a better result than the other two technique when used against a relatively balanced data.

Keywords : recommender system, collaborative filtering, slope one algorithm

1 Pendahuluan

1.1 Latar Belakang

Dengan kemajuan teknologi informasi dan komunikasi, telah mendorong percepatan di segala aspek kehidupan. Dengan memanfaatkan teknologi informasi dan komunikasi maka setiap orang dapat dengan mudah mendapatkan informasi. Seiring berjalannya waktu kebutuhan akan informasi meningkat tetapi tidak sejalan dengan kemampuan jaringan internet. Ledakan informasi (*information overload*) pada jaringan internet menyebabkan seseorang sulit untuk mendapatkan informasi yang cepat dan sesuai kebutuhan.

Untuk itu diperlukan sistem yang dapat mengakses dan menemukan informasi sesuai dengan kebutuhan. Banyak berkembang pencarian menggunakan bantuan mesin pencari (*search engine*), dengan mesin ini pencari informasi telah memiliki referensi yang jelas dan mengetahui dengan pasti informasi apa yang hendak dicari. Sedangkan apa yang terjadi ketika pencari informasi tidak berbekal referensi dan hanya memiliki topik tertentu. Pencarian ini tidak cukup dibantu hanya dengan *search engine* tetapi juga memerlukan suatu sistem penyedia rekomendasi (*recommender system*) dalam menemukan informasi yang diinginkan.

Recommender system merupakan suatu sistem yang merekomendasikan mengenai suatu hal tertentu kepada pengguna sesuai dengan profil pengguna tersebut. *Recommender system* membantu dalam mengatasi masalah *information overload* dengan menyediakan saran-saran bersifat personal berdasarkan pada *history* profil pengguna sebelumnya. Berdasarkan algoritma filtering yang digunakan, *recommender system* dibagi menjadi tiga jenis, yaitu *Demographic Filtering*, *Content-based Filtering*, dan *Collaborative (social) Filtering*. *Recommender system* memiliki beberapa metode dan salah satu metode tersukses pada *recommender system* adalah *collaborative filtering* [6]. Teknik ini akan bekerja dengan cara menghimpun *feedback* pengguna dalam bentuk rating bagi item-item dalam suatu domain yang diberikan dan memanfaatkan kemiripan dan perbedaan antar profil dari beberapa pengguna dalam menentukan bagaimana merekomendasi suatu item selain itu teknik ini juga dapat merekomendasikan item yang relevan kepada pengguna.

Collaborative Filtering memiliki 2 (dua) skema yaitu *memory based* dan *model based*. *Memory based* menggunakan ukuran similarity untuk menghasilkan sebuah prediksi, sedangkan *model based* menggunakan model *user preferences* yang digunakan pada prediksi rating. Dalam skema *model based Collaborative Filtering* penulis menggunakan *Slope One Predictor* karena kelebihan yang ada pada algoritma ini, yaitu mudah diimplementasikan, mudah untuk melakukan *update* dan efisien saat melakukan *query*[5].

Slope One Predictor memiliki 3 (tiga) algoritma yang bisa digunakan yaitu *Slope One*, *Weighted Slope One* dan *Bi-Polar Slope One* [3]. *Slope one predictor* memberikan prediksi rating pada seorang *user* tentang suatu *item* berdasarkan *rating* yang telah diberikan oleh *user* lain pada *item* tersebut. Dari ketiga jenisnya *bi-polar slope one* memiliki kelebihan yaitu tingkat akurasi prediksinya bisa lebih baik dari *slope one* dan *weighted slope one* tergantung teknik pembagian data yang digunakan [3].

Pada algoritma *bi-polar slope one* untuk menentukan prediksi rating data akan dibagi menjadi dua bagian, yaitu data dengan rating yang disukai dan data dengan rating yang kurang disukai. Secara normal pada sebuah skala 1 sampai 10, *item* dengan rating 1 sampai 5 bisa dikatakan kurang disukai, dan rating diatas 5 bisa dikatakan disukai. Hal ini tidak menjadi masalah apabila persebaran data cukup merata, yaitu perbandingan antara rating yang disukai dan kurang disukai cukup berimbang. Masalah bisa timbul bila persebaran data kurang merata sehingga perbandingan rating yang disukai dan kurang disukai sangat tidak berimbang [3].

Untuk mengatasi ketidakseimbangan tersebut, pembagian data bisa dilakukan berdasarkan persebaran data sehingga perbandingan data yang dihasilkan bisa cukup berimbang. Dengan cara ini maka perbandingan data yang disukai dan kurang disukai dapat dibuat cukup berimbang sehingga diharapkan hasil prediksi dapat mencapai nilai yang optimum.

Dalam penyusunan tugas akhir ini dilakukan implementasi algoritma *bi-polar slope one* pada perhitungan prediksi berbasis *collaborative filtering*. Analisis yang dilakukan mengenai pengaruh teknik pembagian data terhadap hasil prediksi menggunakan algoritma *bi-polar slope one*. Diharapkan dengan analisis ini dapat diketahui teknik pembagian data yang dapat menghasilkan hasil prediksi yang optimum.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diungkapkan sebelumnya, maka dapat dirumuskan permasalahan sebagai berikut :

- a. Bagaimana menerapkan *recommender system* berbasis *collaborative filtering* dengan menggunakan algoritma *bi-polar slope one* ?
- b. Bagaimana pengaruh pembagian data terhadap hasil prediksi menggunakan algoritma *bi-polar slope one* ?

Batasan-batasan yang digunakan dalam penyusunan tugas akhir ini sebagai berikut :

- a. *Dataset* yang digunakan bersumber dari *record* MovieLens.
- b. Aplikasi yang akan dibangun berbasis web.
- c. Tidak membahas masalah tingkat kepuasan *user* terhadap hasil prediksi.
- d. Hasil akhir sistem adalah prediksi nilai *rating* dimana hal ini dapat dijadikan pedoman dalam melakukan rekomendasi *item*

1.3 Tujuan

Berdasarkan rumusan masalah di atas, maka tujuan dari tugas akhir ini sebagai berikut :

- a. Mengimplementasikan algoritma *bi-polar slope one* pada *recommender system* berbasis *collaborative filtering*.
- b. Mengukur dan menganalisis pengaruh pembagian data terhadap hasil prediksi menggunakan algoritma *bi-polar slope one* untuk mengetahui teknik yang dapat memberikan hasil prediksi yang optimum.

1.4 Metodologi Penyelesaian Masalah

- a. Studi Literatur
Referensi pada buku, paper serta sumber lain terkait dengan *recommender system*, *collaborative filtering*, dan *slope one predictor* yang digunakan sebagai pedoman dalam penyusunan tugas akhir.
- b. Pengumpulan dan Pengolahan Data
Data yang dikumpulkan merupakan *dataset* yang bersumber dari *record* MovieLens.
- c. Analisis dan Desain
Analisis serta perancangan terhadap *recommender system* yang akan dibangun.
- d. Implementasi
Tahap pembangunan sesuai rancangan yang telah dibangun sebelumnya serta menerapkan algoritma.
- e. Testing
Melakukan pengujian terhadap perangkat lunak yang dibangun. Pengujian dilakukan berdasarkan skenario pengujian yang telah ditentukan sebelumnya.
- f. Analisis Hasil
Melakukan analisis dari implementasi sistem dan pengujian hasil sesuai skenario yang telah ditentukan. Mean Absolute Error (MAE) digunakan untuk menghitung akurasi dari prediksi yang dihasilkan.
- g. Pengambilan kesimpulan dan penyusunan laporan tugas akhir
Mengambil kesimpulan dari analisis hasil yang telah dilakukan, kemudian menyusun laporan dari analisis hasil yang telah dilakukan.

1.5 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, metodologi penyelesaian masalah, dan sistematika penulisan.

BAB II LANDASAN TEORI

Berisi uraian mengenai landasan teori yang akan digunakan, meliputi teori tentang algoritma *slope one* dan teori-teori lain yang berkaitan dengan penelitian tugas akhir ini.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Berisi tentang analisa dan perancangan terhadap *recommender system* yang akan dibangun.

BAB IV ANALISIS DAN PENGUJIAN SISTEM

Berisi implementasi dari hasil analisa dan perancangan sistem yang dibuat, serta pengujian sistem.

BAB V KESIMPULAN DAN SARAN

Berisi kesimpulan dan saran-saran untuk pengembangan lebih lanjut terhadap hasil penelitian ini.

5 Penutup

5.1 Kesimpulan

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

1. Pada algoritma *bi-polar slope one*, teknik pembagian data lebih berpengaruh dibanding jumlah *training set* dalam membuat prediksi *rating*. Makin seimbang pembagian data yang dilakukan, hasil prediksinya semakin baik.
2. Pada data dengan persebaran yang merata, pengukuran menggunakan MAE pada teknik *auto balance* selalu lebih baik dibanding teknik pembagian data dengan *lainnya*.
3. Pada pengukuran akurasi kecocokan hasil prediksi, pembagian data dengan *auto balance* menunjukkan peningkatan hasil yang signifikan seiring bertambahnya jumlah *training set*. Tidak menutup kemungkinan pada jumlah data yang lebih banyak *teknik auto balance* akan menghasilkan kecocokan hasil prediksi lebih baik dari dua teknik lainnya.
4. Nilai MAE yang dihasilkan oleh sistem masih berkisar pada angka 0,7 hingga 0,9 dengan tingkat akurasi sebesar 30 hingga 40 persen. Hal ini menunjukkan bahwa tingkat akurasi sistem pada level nilai *rating* masih kurang baik.

5.2 Saran

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

1. Untuk membuktikan bahwa persentase kesamaan *rating* teknik *auto balance* bisa lebih baik maka bisa dilakukan pengujian pada jumlah *training set* yang lebih banyak.
2. Membandingkan secara *head-to-head* seluruh keluarga algoritma *slope one* menggunakan skenario pengujian yang sama.
3. Mengimplementasikan teknik pembagian data di atas pada data yang tidak seimbang secara ekstrim untuk mengetahui apakah hasilnya akan sama dengan pada data yang seimbang.
4. Mengimplementasikan algoritma *bi-polar slope one* pada suatu *recommender system* untuk memberikan rekomendasi item bagi user.

6 Daftar Pustaka

- [1] Deshpande, M., Karypis, G., Item-based Top-N Recommendation Algorithms. 2003.
- [2] Eppler, M. J. and Mengis, Jeanne . A Framework for Information Overload Research in Organizations. Svizzera University . Italia. 2003.
- [3] Lemire, D., Maclachlan, A. Slope One Predictors for Online Rating-Based Collaborative Filtering. 2005.
- [4] Lemire, D., McGrath, S. Implementing Rating-Based Item-to-Item Recommender System in PHP/SQL. 2005.
- [5] Rani, T.P. Analisis dan Implementasi Algoritma Slope One pada Recommender System berbasis Collaborative Filtering. 2009.
- [6] Sarwar, B., Karypis, G., Konstan, J., Riedl, J. Item-based Collaborative Filtering Recommendation Algorithm. 2001.
- [7] Vozalis, E. Margaritis, K. G. Analysis of Recommender Systems Algorithms. 2003.