

Uji Akurasi Algoritma Bipolar Slope One Dan BW-Mine Pada Sistem Rekomendasi

Ali Akbar Lubis¹, Ronsen Purba², Megawaty Simamora³, Anna Agustiana⁴

Jurusan Teknik Informatika, STMIK Mikroskil

¹ali.akbar@mikroskil.ac.id, ²ronsen@mikroskil.ac.id, ³141114762@students.mikroskil.ac.id,

⁴141113147@students.mikroskil.ac.id

Abstrak

Sistem rekomendasi banyak diterapkan pada berbagai e-commerce. Ada berapa masalah yang dapat menyebabkan sistem rekomendasi untuk gagal. Masalah ini adalah masalah kekosongan data rating yang massif (sparsity) dan cold start. Oleh karena itu, diperlukan metode rekomendasi yang tepat untuk meningkatkan keakurasian, sehingga user dapat menemukan item yang sesuai dengan keinginannya.

Untuk mencapai tujuan tersebut, digunakan bipolar slope one untuk memprediksi rating. Bipolar slope one digunakan untuk memprediksi rating suatu item. Dalam memprediksi rating suatu item, dibutuhkan pola item. Pola item ini dapat direpresentasikan dalam Assosiation Rule yang terdapat pada algoritma BW-Mine.

Pengujian dilakukan dengan MAE yang melibatkan 50 user terhadap 200 item. Hasil pengujian menggunakan MAE diperoleh bahwa sparsity memiliki pengaruh terhadap akurasi prediksi rating yang dihasilkan dalam sistem rekomendasi.

Kata Kunci : Sistem Rekomendasi, Bipolar slope one, BW-Mine,

Abstract

The recommendation system is widely applied to various e-commerce. There are some problems that can cause the recommendation system to fail. This problem is about the massive vacuum of rating data (sparsity) and cold start. Therefore, the right recommendation method is needed to improve accuracy, so that the user can find the item according to desire.

To achieve this goal, bipolar slope one is used to predict the rating. Bipolar slope one is used to predict the rating of an item. In predicting an item's rating, an item pattern is needed. This item pattern can be represented in the Assosiation Rule that found in the BW-Mine algorithm.

The test was carried out with MAE involving 50 users of 200 items. The test results using MAE, obtained that sparsity has an influence on the accuracy of rating prediction generated in the recommendation system

Keyword : recommendation system, Bipolar Slope One, BW-Mine

1. PENDAHULUAN

Sistem rekomendasi merupakan sistem berbasis pengetahuan yang mencoba memprediksi produk atau layanan yang paling cocok berdasarkan preferensi user [1]. Tapi ada berapa masalah yang dapat menyebabkan sistem rekomendasi untuk gagal. Masalah ini adalah masalah kekosongan data rating yang massif (sparsity) dan cold start user ialah seorang user yang baru bergabung dengan suatu aplikasi dan belum memberikan rating untuk item apapun [2]. Sparsity data rating akan memberikan poor prediction terhadap hasil rekomendasi [3]. Oleh

karena itu, diperlukan metode rekomendasi yang tepat untuk meningkatkan keakurasian, sehingga user dapat menemukan item (film) yang sesuai dengan keinginannya untuk ditonton.

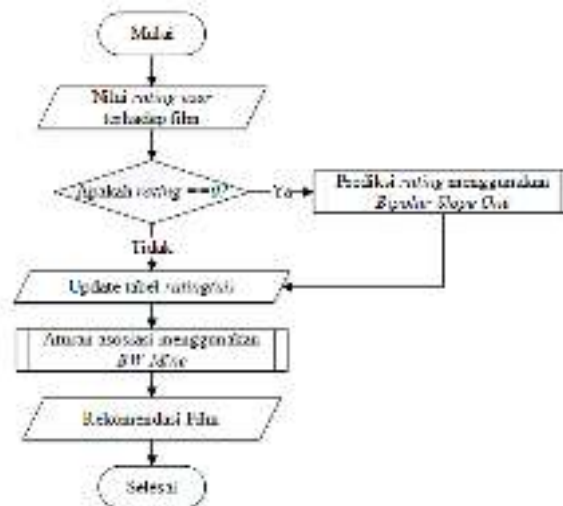
Salah satu cara mengatasi sparsity dengan melakukan prediksi rating menggunakan algoritma Bipolar Slope One. Algoritma ini memprediksi rating suatu item dengan cara memisahkan menjadi dua bagian yaitu items yang disukai user dan items yang tidak disukai user [4]. Langkah pertama untuk item, melakukan devisiasi antara dua item yang disukai atau devisiasi antara dua item yang tidak disukai [4]. Langkah kedua untuk user, hanya devisiasi dari pasangan user yang memberikan rating pada item i dan item j , dimana item i digunakan untuk memprediksi rating item j . Prediksi rating digunakan untuk mengisi rating yang masih kosong, sehingga permasalahan sparsity dapat teratasi.

Dalam membantu user baru atau user yang belum memberikan rating untuk item diperlukan informasi mengenai pola film. Pola ini didapat dengan menemukan hubungan penting di antara film yang telah ditonton oleh user. Hubungan ini direpresentasikan dalam bentuk Association Rule atau aturan item yang sering muncul [5], salah satu algoritma BW-Mine. BW-Mine merupakan algoritma yang dijalankan secara searah, yang bekerja dengan dataset padat. Algoritma ini menampung pola yang sering dan kemudian membentuk Association Rule Mining searah sehingga algoritma ini sangat cocok untuk dataset yang besar [6]. Manfaat dari Association Rule Mining ini mendeteksi hubungan yang tidak diketahui, memberikan hasil yang dapat membantu user mendapatkan informasi pola item yang saling berelasi

2. METODE PENELITIAN

2.1 Analisis Proses

Analisis proses adalah pendeskripsian dan pengidentifikasian proses berjalannya program yang dibangun. Gambar 1 menunjukkan proses kerja sistem rekomendasi film dengan algoritma Bipolar Slope One dan BW – Mine menggunakan flowchart.

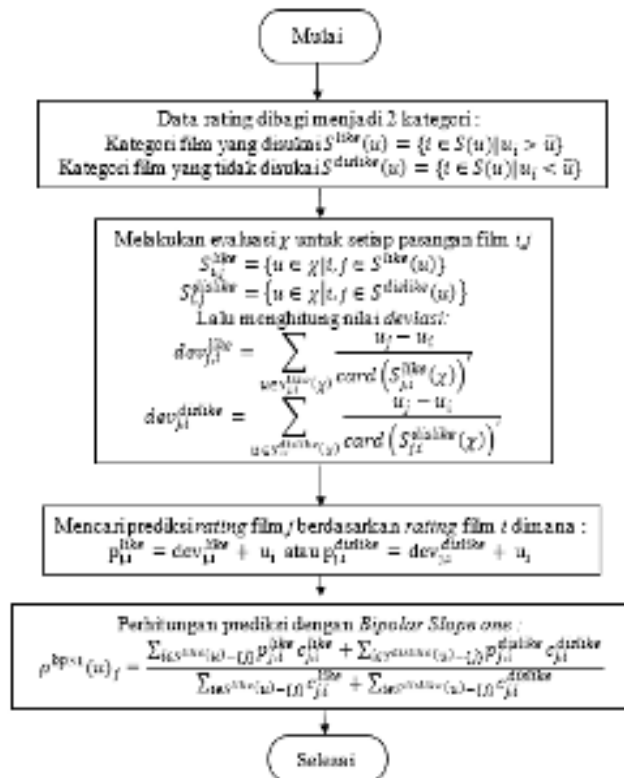


Gambar 1 Flowchart Sistem Rekomendasi Film

2.2 Proses Prediksi Rating Menggunakan Bipolar Slope One

Algoritma Bipolar Slope One memprediksi rating berdasarkan informasi dari user lain. Pada kasus ini algoritma tidak memperhatikan jenis film, tetapi memperhatikan kemiripan pola

rating sesama user. Berikut proses kerja sistem rekomendasi film dengan algoritma Bipolar Slope One menggunakan flowchart pada gambar 2.



Gambar 2 Flowchart Algoritma Bipolar Slope One

Cara kerja algoritma bipolar slope one yaitu membagi prediksi menjadi dua bagian yaitu prediksi item yang disukai user dan prediksi item yang tidak disukai user. Untuk mendukung semua jenis user maka diterapkan rata – rata user sebagai ambang batas antara items yang disukai dan yang tidak disukai oleh user. Bipolar Slope One membatasi rating bersifat prediksi. Langkah pertama untuk item, hanya membagi dua antara dua item yang disukai atau membagi dua antara dua item yang tidak disukai. Sehingga rating item dapat dirumuskan menjadi 2 yaitu [4]:

$$S_{i,j}^{like} = \{u \in \chi | i, j \in S^{like}(u)\} \quad (1)$$

$$S_{i,j}^{dislike} = \{u \in \chi | i, j \in S^{dislike}(u)\} \quad (2)$$

Keterangan:

$S_{i,j}^{like}$ = rating item yang disukai antara i dan j

$S_{i,j}^{dislike}$ = rating item yang tidak disukai antara i dan j

Menggunakan dua rangkaian tersebut maka akan dihitung matriks deviasi untuk item yang disukai dan matriks deviasi untuk item yang tidak disukai dengan menggunakan:

$$dev_{j,i}^{like} = \sum_{u \in S_{j,i}^{like}(X)} \frac{u_j - u_i}{card(S_{j,i}^{like}(X))} \quad (3)$$

Keterangan:

$dev_{j,i}^{like}$ = Rata – rata selisih rating item i dan item j yang disukai oleh user.

u = User yang memberikan rating item i dan item j

u_j = Rating untuk item j

u_i = Rating untuk item i
 x = Himpunan rating items dari user.
 $S_{j,i}^{like}(\chi)$ = Himpunan bagian dari x yang mengandung item i dan item j yang disukai oleh user
 $card(S_{j,i}^{like}(\chi))'$ = Jumlah user yang memberikan rating item i dan item j (jumlah elemen di $S_{j,i}(\chi)$)

$$dev_{j,i}^{dislike} = \sum_{u \in S_{j,i}^{dislike}(\chi)} \frac{u_j - u_i}{card(S_{j,i}^{dislike}(\chi))'} \quad (4)$$

Keterangan:

$dev_{j,i}^{dislike}$ = Rata – rata selisih rating item i dan item j yang tidak disukai oleh user.
 $S_{j,i}^{dislike}(\chi)$ = Himpunan bagian dari x yang mengandung item i dan item j yang tidak disukai oleh user

$card(S_{j,i}^{dislike}(\chi))'$ = Jumlah user yang memberikan rating item i dan item j (jumlah elemen di $S_{j,i}(\chi)$)

Prediksi untuk rating item j berdasarkan rating item i dimana adalah $p_{j,i}^{like} = dev_{j,i}^{like} + u_i$ atau $p_{j,i}^{dislike} = dev_{j,i}^{dislike} + u_i$ bergantung pada masing - masing terdapat di $S^{like}(u)$ atau $S^{dislike}(u)$. Maka prediksi *Bipolar Slope One* dihitung menggunakan persamaan dibawah ini, dimana bobot $c_{j,i}^{like} = card(S_{j,i}^{like})$ dan $c_{j,i}^{dislike} = card(S_{j,i}^{dislike})$ [4].

$$\rho^{bpS1}(u)_j = \frac{\sum_{i \in S^{like}(u) - \{j\}} p_{j,i}^{like} c_{j,i}^{like} + \sum_{i \in S^{dislike}(u) - \{j\}} p_{j,i}^{dislike} c_{j,i}^{dislike}}{\sum_{i \in S^{like}(u) - \{j\}} c_{j,i}^{like} + \sum_{i \in S^{dislike}(u) - \{j\}} c_{j,i}^{dislike}} \quad (5)$$

Keterangan :

$\rho^{bpS1}(u)_j$ = Prediksi *Bipolar Slope One* terhadap item j .

$dev_{j,i}^{like}$ = Rata-rata rating item i terhadap item j

$p_{j,i}^{like}$ = Prediksi terhadap item j berdasarkan item i pada kategori item yang disukai.

$p_{j,i}^{dislike}$ = Prediksi terhadap item j berdasarkan item i pada kategori item yang tidak disukai.

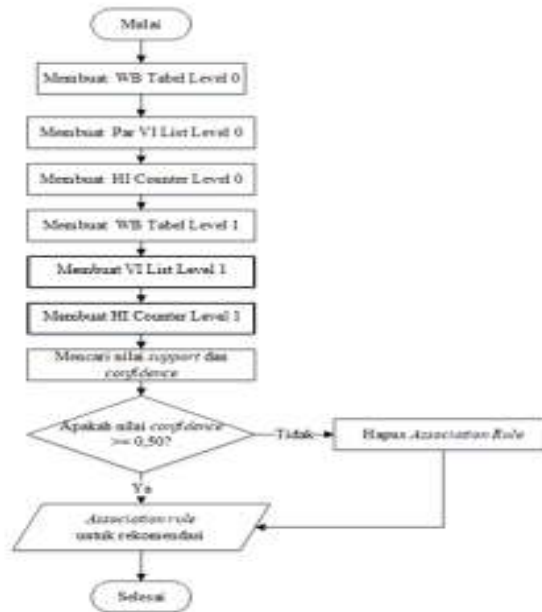
$c_{j,i}^{like}$ = Jumlah user yang memberikan rating kedua item yang dibandingkan pada kategori item yang disukai.

$c_{j,i}^{dislike}$ = Jumlah user yang memberikan rating kedua item yang dibandingkan pada kategori item yang tidak disukai.

2.3 Proses Pembentukan Aturan Asosiasi Menggunakan BW-Mine

BW-Mine (bitwise mining) adalah metode yang menggunakan history dari setiap user untuk merekomendasikan item yang sesuai dan relevan kepada user [6] Algoritma BW-Mine memindai dataset hanya sekali dan membuat beberapa tabel. Mine dalam data besar digunakan untuk menemukan informasi yang berguna dan pengetahuan berharga tentang pola akses, algoritma ini dapat menghasilkan sering item yang ditetapkan secara paralel sehingga efisien untuk masalah yang memiliki data besar juga. Menemukan pola yang sering dari sistem rekomendasi harus memiliki konten penting dari items yang terkait pada bitwise table (WB-table). Pembentukan aturan asosiasi pada algoritma BW – Mine digunakan untuk menemukan

hubungan antar items berdasarkan nilai rating yang sudah ada atau nilai rating yang diperoleh dari algoritma Bipolar Slope One. Berikut gambar 3 yang merupakan flowchart proses pembentukan aturan asosiasi menggunakan algoritma BW – Mine.



Gambar 3 Flowchart Algoritama *BW-Mine*

Cara kerja Algoritma BW-Mine adalah [10]:

1. Langkah pertama menentukan minsup are frequent, lalu membentuk struktur data WB – Table Level 0 dengan melakukan pemindai rating user, mengidentifikasi item yang menarik dan item yang tidak menarik untuk user.
2. Untuk setiap baris dari WB-table Level 0 menciptakan ParVI-List (Parallel-Computation Based Vertical Index List) Level 0 yang merupakan struktur data dua dimensi. ParVI-List Level 0 hanya digunakan untuk membentuk WB – Table Level1 [10].
3. Struktur data berikutnya yang digunakan adalah Level-0 HI-Counter yang menyimpan counter setiap Item. Untuk setiap kolom c dari WB-table, kita menciptakan Level-0 HI-Counter dengan menghitung jumlah 1s dalam kolom c dan menempatkan menghitung dalam posisi c-th dari HI-Counter [10].
4. Selanjutnya, algoritma BW-Mine menemukan Single Frequent item dengan memeriksa Level-0 HI-Counter. Lalu menentukan minsup are frequent, lalu item dengan frequent \geq minsup are frequent dilanjutkan untuk langkah berikutnya.
5. Langkah berikutnya adalah membentuk WB – Table Level 1, lalu membentuk VI-List Level 1 dengan merekam indeks kolom yang dilakukan dalam sekali pemindaian pada WB – Table Level 1. Dimana column index merupakan items dan row indices merupakan user tertarik pada item tersebut. Semua langkah dapat diselesaikan hanya dalam satu pemindaian pada VI-List karena setiap user hanya
6. Dari setiap r baris dari VI- List, BW-Mine mempertimbangkan baris masing-masing di WB-table untuk mendapatkan database untuk X (Item) dan dengan cara yang sama level-1 HI-Counter dibuat sampai level seterusnya hingga mendapatkan semua pola yang dihasilkan.
7. Mencari nilai confidence untuk semua pola yang membentuk association rule. Association rule diurutkan secara descending berdasarkan nilai confidence:

$$\text{confidence}(A, B) = \frac{\text{visited } A \cup B}{\text{visited } A} \quad (6)$$

Keterangan:

$\text{visited } A \cup B$ = total frekuensi yang mengunjungi *item* A dan *item* B

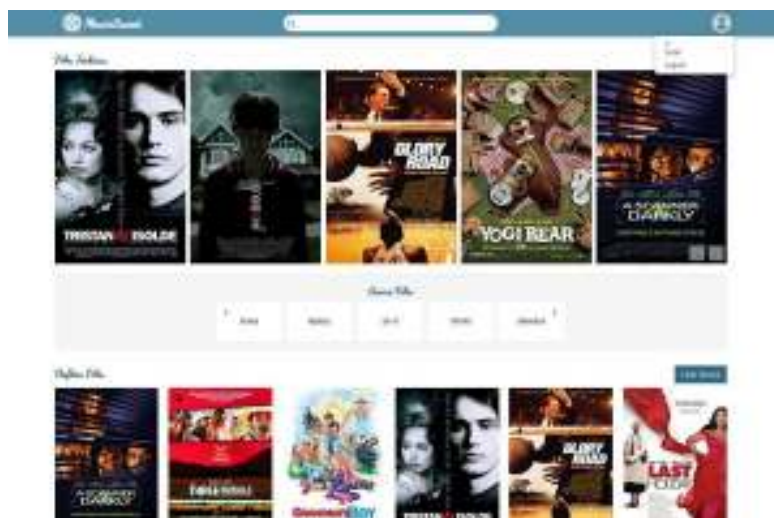
$\text{visited } A$ = total frekuensi yang mengunjungi *item* A

8. Mencari *frequent itemsets* dari transaksi pada database sudah ditemukan, maka langkah selanjutnya adalah menemukan aturan asosiasi yang kuat (*strong association rule*) dari *frequent itemsets*. Aturan asosiasi yang kuat memenuhi syarat dari *minimum support* dan *minimum confidence* akan direkomendasikan kepada *user*.

3. HASIL DAN PEMBAHASAN

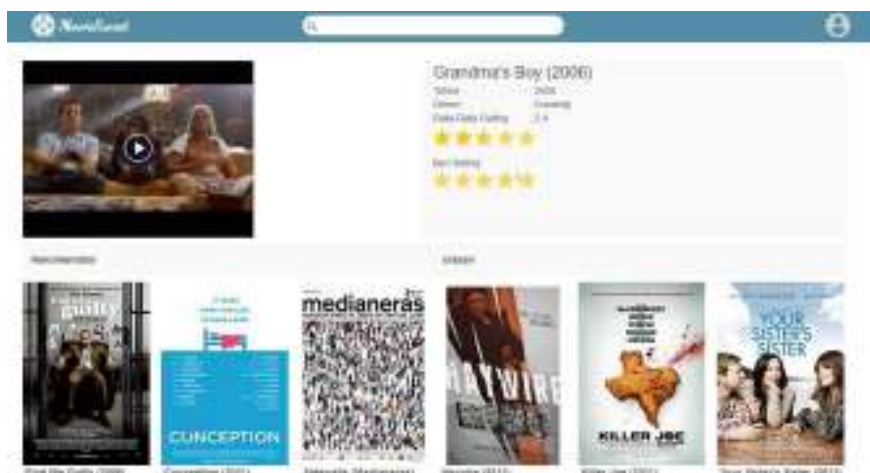
3.1 Hasil

Berikut tampilan dari Sistem rekomendasi film yang terlihat pada gambar 4:



Gambar 4 Tampilan Halaman Home

Gambar 4 adalah tampilan halaman *Home* untuk *member* yang telah melakukan *login*. *Member* dapat melihat terdapat *searchBox*, *link Daftar*, *link Login*, film terbaru, daftar film dan Menu *list genre* yang dapat dijadikan *filter* film berdasarkan *genre*.



Gambar 5 Halaman Detail Film untuk member

Gambar 5 merupakan tampilan halaman detail film untuk *member*. Adapun selain terdapat informasi film berupa judul, tahun, *genre*, rata-rata *rating*, *member* juga dapat memutar *trailer* dari film dan memberi *rating*. Hasil rekomendasi terkait dengan item yang dipilih dapat dilihat pada daftar film di bagian bawah.

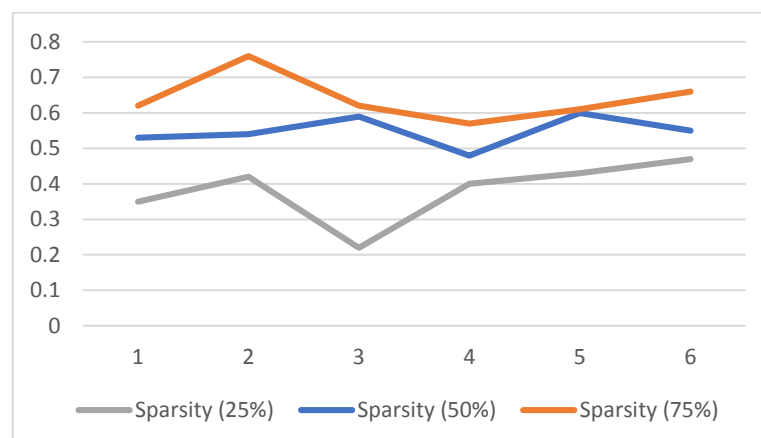
3.2 Pengujian

Pengujian algoritma *bipolar slope one* dilakukan dengan *Mean Absolute Error* (MAE) untuk mengukur keakuratan hasil prediksi *rating* berdasarkan tingkat kekosongan data (*sparsity*). Data sample yang digunakan sejumlah nilai *rating* yang diberikan oleh 50 *user* terhadap 200 film. Pengujian dilakukan dengan menghitung nilai error MAE berdasarkan tingkat kekosongan data *rating* 25%, 50%, dan 75%. Pengujian ini untuk mengetahui pengaruh *sparsity* data terhadap akurasi prediksi *rating* yang dihasilkan oleh algoritma *bipolar slope one*. Nilai MAE diperoleh dengan mengurangkan nilai *rating* sebenarnya dengan nilai prediksi *rating*, dimana nilai *rating* sebenarnya diperoleh dari rata-rata *rating* film terhadap semua *user*.

Tabel 1 Pengujian hasil MAE algoritma *Bipolar Slope One*

Percobaan	MAE		
	Sparsity (25%)	Sparsity (50%)	Sparsity (75%)
1	0.35	0.53	0.62
2	0.42	0.54	0.76
3	0.22	0.59	0.62
4	0.40	0.48	0.57
5	0.43	0.60	0.61
6	0.47	0.55	0.66

Pada table 1 dapat dilihat bahwa pengujian yang dilakukan dengan 6 percobaan yang masing-masing dilakukan dengan tingkat *sparsity* yang berbeda-beda (25%, 50%, dan 75%). Dan hasil dari pengujian tersebut terlihat semakin kecil tingkat *sparsity* maka nilai MAE akan semakin kecil. Ini artinya semakin kecil tingkat *sparsity* maka akurasi akan semakin baik dalam menghasilkan prediksi *rating*. Sebaliknya semakin besar tingkat *sparsity* maka nilai MAE akan semakin besar pula atau nilai akurasi semakin kecil. Untuk lebih jelasnya dapat dilihat pada gambar 6 di bawah ini.



Gambar 6 Grafik perbandingan MAE terhadap tingkat *sparsity*

Pada gambar 6 di atas dapat dilihat bahwa nilai MAE dengan sparsity 75% memiliki kecenderungan berada di atas nilai MAE dengan sparsity 50%. Begitupun untuk nilai MAE dengan sparsity 50% memiliki kecenderungan berada di atas nilai MAE dengan sparsity 25%. Oleh karena itu, dapat dikatakan nilai sparsity dapat mempengaruhi nilai akurasi dari prediksi rating untuk sistem rekomendasi

4. KESIMPULAN

Berdasarkan penerapan algoritma Bipolar Slope One dan BW-Mine pada Sistem Rekomendasi yang dibangun, maka dapat disimpulkan tingkat kekosongan data rating (sparsity) mempengaruhi nilai MAE, MAE tertinggi adalah 0.76 dengan tingkat kekosongan data 75%. Dan Nilai MAE terendah terdapat pada tingkat kekosongan data 25% dengan nilai MAE adalah 0.22.

5. SARAN

Adapun saran yang dapat diberikan dan membantu dalam penelitian selanjutnya adalah Sebaiknya ditambahkan algoritma yang melakukan pendekatan terhadap kemiripan user agar item yang belum pernah dirating juga bisa direkomendasikan kepada user seperti algoritma Trust-Aware Recommendation.

DAFTAR PUSTAKA

- [1] Ricci, F., Rokach, L. & Shapira, B., 2011. Introduction Recommender System Handbook. Dalam: Recommender Systems Handbook. New York: Springer.
- [2] Yang, B., Lei, Y., Lui, D. & Lui, J., 2016. Social Collaborative Filtering by Trust. Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, pp. 2747-2753 tersedia pada: <https://www.ijcai.org/Proceedings/13/Papers/404.pdf>, tanggal akses 26 Desember 2018
- [3] Li, B., Yang, Q. & Xue, X., 2009. Can Movies and Book Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction. IJCAI. tersedia pada: <https://www.ijcai.org/Proceedings/09/Papers/338.pdf>, tanggal akses 18 November 2018
- [4] Lemire, D. & Maclachlan, A., 2005. Slope One Predictors for Online Rating-Based Collaborative Filtering..
- [5] Prasetyo, E., 2012. Data Mining: Konsep dan Aplikasi Menggunakan Matlab. Yogyakarta: Andi
- [6] Makkar, P. & Shintre, S., 2018. Recommendation Based On Comparative Analysis of Apriori and BW-Mine Algorithm. International Journal of Advanced Engineering, Management and Science (IJAEMS), 4(2), pp. 95-99
- [7] Gurfoni, A. I. & Sulastri, H., 2017. Penerapan Data Mining dalam Pengelompokan Penderita Thalassaemia. Jurnal Teknologi dan Sistem Informasi, Volume 3, pp. 299-305.
- [8] Xiao, B. & Izak, B., 2007. E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact. MIS Quarterly, 31(1), pp. 137-209.
- [9] Sarwar, b., Karypis, G., Konstan, J. & Riedl, J., 2001. Item-based collaborative filtering recommendation algorithm. tersedia pada: Available at: <http://www.ra.ethz.ch/cdstore/www10/papers/pdf/p519.pdf>, tanggal akses 7 November 2018
- [10] Leung, C. K., Jiang, F. & Pazdor, A. G. M., 2017. Bitwise Parallel Association Rule Mining for Web Page Recommendation. Association for Computing Machinery, pp. 662-669.