

## SISTEM REKOMENDASI: BUKU ONLINE DENGAN METODE COLLABORATIVE FILTERING

Moh. Irfan<sup>1</sup>, Andharini Dwi C<sup>2</sup>, Fika Hastarita R.<sup>3</sup>

<sup>1,2,3</sup> Program Studi Teknik Informatika, Universitas Trunojoyo Madura

Masuk: 5 April 2014, revisi masuk : 19 Juni 2014, diterima: 7 Juli 2014

### ABSTRACT

The book is a source of information regarding all aspects of life, especially education. However, low interest in reading among the public is a major issue in education today. Recommendation systems can help recommend the reader to more easily obtain information about the books to be read. Therefore, in this study made an online book recommendation system using Collaborative Filtering. Collaborative Filtering is one of the methods that can be used in making the recommendation system. The results of this study showed that the average value of the MAE (Mean Absolute Error) on trial 1 (1.064) is smaller than 2 trials (1.21), 4 trials (2,474) and test 5 (3.526). This shows that the more the amount of data used and if there is a user who has never rate a, then the resulting system is relatively inaccurate and generate recommendations if using Collaborative Filtering bad.

**Keywords:** recommendation system, Collaborative Filtering, Online Book.

### INTISARI

Buku merupakan sumber informasi semua aspek kehidupan khususnya pendidikan. Namun rendahnya minat baca dikalangan masyarakat menjadi persoalan penting di dunia pendidikan saat ini. Sistem rekomendasi dapat membantu merekomendasikan para pembaca agar lebih mudah mendapatkan informasi mengenai buku yang akan dibaca. Oleh karena itu, pada penelitian ini dibuat sistem rekomendasi buku online menggunakan metode *Collaborative Filtering*. *Collaborative Filtering* adalah salah satu metode yang dapat digunakan dalam membuat sistem rekomendasi. Hasil dari penelitian ini menunjukkan bahwa rata-rata nilai MAE (*Mean Absolute Error*) pada uji coba 1 (1,064) lebih kecil daripada uji coba 2 (1,21), uji coba 4 (2,474) dan ujicoba 5 (3,526). Hal ini menunjukkan bahwa semakin banyak jumlah data yang digunakan dan jika terdapat user yang belum pernah merating, maka sistem yang dihasilkan relatif tidak akurat dan menghasilkan rekomendasi yang buruk jika menggunakan *Collaborative Filtering*.

**Kata Kunci :** Sistem Rekomendasi, *Collaborative Filtering*, Buku Online

### PENDAHULUAN

Buku merupakan informasi segala kebutuhan yang diperlukan, dimulai dari iptek, seni budaya, ekonomi, politik, sosial dan pertahanan keamanan dan lain-lain. Upaya membaca buku membuka wawasan dunia intelek sehingga dapat mengubah masa depan serta mencerdaskan akal, pikiran dan iman. Dengan membaca buku, selain pengetahuan akan semakin bertambah, pribadi akan semakin kaya, yang kesemuannya jelas akan menurunkan

efek negatif terhadap anak-anak, yakni kenakalan. Sedangkan anak yang tidak terbina minat bacanya sejak dini akan menghadapi peluang yang semakin kecil untuk mengembangkan pengetahuan setinggi-tingginya. Namun berdasarkan laporan Bank Dunia, Indonesia merupakan negara yang memiliki minat baca sangat rendah. Hal tersebut sungguh disayangkan, mengingat sebagai negara besar, Indonesia memiliki potensi besar untuk menjadi negara yang unggul.

<sup>1</sup> [irfan09017@gmail.com](mailto:irfan09017@gmail.com)

<sup>2</sup> [andharini.dwi.cahyani@gmail.com](mailto:andharini.dwi.cahyani@gmail.com)

Rendahnya minat baca di kalangan masyarakat menjadi persoalan penting di dunia pendidikan saat ini. Untuk itu diperlukan sebuah sistem yang dapat membantu merekomendasikan para pembaca agar lebih mudah mendapatkan informasi buku-buku yang akan dibaca selanjutnya.

Sistem rekomendasi sendiri telah digunakan secara luas oleh hampir semua area bisnis dimana seorang konsumen memerlukan informasi untuk membuat suatu keputusan. Terdapat dua pendekatan yang umumnya digunakan dalam membuat sistem rekomendasi, yaitu *content based filtering* dan *collaborative filtering*. *Content based filtering* merupakan metode yang bekerja dengan mencari kedekatan suatu item yang akan direkomendasikan ke *user* dengan *items* yang telah diambil oleh pengguna sebelumnya berdasarkan kemiripan antar kontennya. Namun, sistem rekomendasi berbasis konten ini masih memiliki kelemahan, yaitu karena semua informasi dipilih dan direkomendasikan berdasarkan konten, maka pengguna tidak mendapatkan rekomendasi pada jenis konten yang berbeda. Selain itu, sistem rekomendasi ini kurang efektif untuk pengguna pemula, karena pengguna yang masih pemula tidak mendapat masukan dari pengguna sebelumnya. (Li, 2002)

Pendekatan lain untuk menutupi kelemahan dari *content based filtering* dikembangkan, yaitu *collaborative filtering*. Sistem *collaborative filtering* adalah metode yang digunakan untuk memprediksi kegunaan item berdasarkan penilaian pengguna sebelumnya.

*Collaborative Filtering* dapat digunakan untuk membuat sistem rekomendasi, akan tetapi perhitungan dalam algoritma sangat bergantung pada hasil rekomendasi. Seperti halnya skenario yang digunakan dalam perhitungan similarity, antara metode pearson correlation dan adjusted cosine similarity memberikan hasil yang berbeda. Berdasarkan beberapa kelebihan dari metode *collaborative filtering*, pada penelitian ini metode ini diterapkan pada pembuatan sistem rekomendasi buku online menggunakan

dataset buku book crossing dengan dilihat akurasi menggunakan beberapa skenario, yaitu dengan menggunakan cold start problem dan non-cold start problem pada perhitungan prediksinya.

## METODE

Sistem rekomendasi merupakan sebuah (web) alat personalisasi yang menyediakan pengguna sebuah informasi daftar item-item yang sesuai dengan keinginan masing-masing pengguna. Sistem rekomendasi menyimpulkan preferensi pengguna dengan menganalisis ketersediaan data pengguna, informasi tentang pengguna dan lingkungannya. Oleh karena itu sistem rekomendasi akan menawarkan kemungkinan dari penyaringan informasi personal sehingga hanya informasi yang sesuai dengan kebutuhan dan preferensi pengguna yang akan ditampilkan di sistem dengan menggunakan sebuah teknik atau model rekomendasi.

Ada beberapa metode atau teknik yang digunakan dalam sistem rekomendasi. Setiap metode disesuaikan dengan permasalahan dalam menghasilkan sebuah informasi yang sesuai. Metode atau pendekatan yang dipilih pada sistem rekomendasi bergantung pada permasalahan yang akan diselesaikan, teknik rekomendasi yang berbeda-beda digunakan untuk aplikasi yang berbeda, dasar dari suatu tujuan dan objektif dari sebuah aplikasi. Dari penelitian terbaru metode atau teknik rekomendasi memiliki beberapa sejumlah kemungkinan klasifikasi. (Uyun, 2011)

Sistem *collaborative filtering* adalah metode yang digunakan untuk memprediksi kegunaan item berdasarkan penilaian pengguna sebelumnya, misalnya cara pemberian *rating* terhadap suatu *item* (Lam, 2004). Metode ini merekomendasikan item-item yang dipilih oleh pengguna lain dengan kemiripan model item dari pengguna saat ini. Walaupun dalam beberapa riset *collaborative filtering* terbukti dapat menutupi beberapa kekurangan pendekatan *content based* dan banyak

diimplementasikan dalam aplikasi nyata, namun pendekatan ini memiliki beberapa kekurangan, antara lain: (Uyun, 2011)

*Cold-start problem*, karena pendekatan *collaborative filtering* melakukan prediksi berdasarkan *rating* yang diberikan user pada *item*, maka menjadi suatu masalah ketika suatu item baru masuk ke dalam sistem dan belum di-*rating* sama sekali oleh user. Akibatnya item tersebut tidak akan pernah direkomendasikan kepada user.

*Sparsity*, untuk ukuran data yang besar, banyak *item* yang baru sedikit di-*rating* oleh user, akibatnya *item* tersebut memiliki nilai prediksi yang relatif tidak akurat dan menghasilkan rekomendasi yang buruk.

Salah satu metode sistem rekomendasi adalah *collaborative filtering*. Berikut ini adalah tahap-tahap memberikan rekomendasi menggunakan *collaborative filtering*.

Dasar perhitungan *similarity* pada *item-based collaborative filtering* antara dua buah *item*  $i$  dan  $j$  adalah dengan mencari user mana saja yang telah memberi *rating* pada *item*  $i$  dan  $j$  lalu gunakan metode perhitungan *similarity*. Pada ICHM terdapat dua buah matriks, matriks *group-rating* dan matriks *item-rating*, maka perhitungan *similarity* juga dilakukan untuk masing-masing matriks lalu hasilnya digabungkan untuk perhitungan prediksi.

Metode *pearson correlation-based similarity* merupakan metode perhitungan berbasis korelasi yang paling banyak diimplementasikan untuk perhitungan nilai *similarity*. Korelasi Pearson mengukur seberapa besar hubungan linear antara dua variabel. Koefisien korelasi Pearson berasal dari model regresi linier yang memiliki asumsi yaitu bahwa hubungan antara dua variabel harus linier, dengan kesalahan harus independen dan memiliki distribusi probabilitas dengan mean 0 dan varians (berdistribusi Normal (0,1)). (Li, 2002) Metode *pearson correlation-based similarity* ditunjukkan oleh Persamaan

$$sim(k,l) = \frac{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)(R_{u,l} - \bar{R}_l)}{\sqrt{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)^2} \sqrt{\sum_{u=1}^m (R_{u,l} - \bar{R}_l)^2}}$$

Keterangan:

$sim(k,l)$  adalah nilai *similarity* antara *item*  $k$  dan *item*  $l$

$m$  adalah jumlah total *user* yang merating *item*  $k$  dan *item*  $l$

$\bar{R}_k$  dan  $\bar{R}_l$  adalah *rating* rata-rata pada *item*  $k$  dan *item*  $l$

$R_{u,k}$  dan  $R_{u,l}$  adalah *rating* yang diberikan oleh *user*  $u$  kepada *item*  $k$  dan *item*  $l$

*Adjust Cosine Similarity*. *Cosine similarity* merupakan metode yang sering digunakan untuk menghitung kesamaan pengguna, tetapi metode ini memiliki satu kekurangan. Perbedaan skala *rating* antara berbagai pengguna akan menghasilkan *similarity* yang sangat berbeda. Sebagai contoh, user A merating buku terbaik dengan *rating* 4 dan tidak pernah member *rating* 5 pada buku apapun, dan member *rating* 1 pada buku terjelek, tidak sesuai dengan tingkat standar *rating* yaitu 2. Tetapi user B selalu merating sesuai dengan tingkat standar, member *rating* 5 pada buku terbaik, dan 2 pada buku yang jelek. Jika menggunakan *cosine similarity*, keduanya sangat berbeda. *Adjusted cosine similarity* mengatasi kelemahan dari *cosine similarity*. (Djamal, 2010)

Metode *Cosine similarity* dapat ditunjukkan oleh Persamaan

$$sim(k,l) = \frac{\sum_{u=1}^m (R_{u,k} - \bar{R}_u)(R_{u,l} - \bar{R}_u)}{\sqrt{\sum_{u=1}^m (R_{u,k} - \bar{R}_u)^2} \sqrt{\sum_{u=1}^m (R_{u,l} - \bar{R}_u)^2}}$$

Keterangan:

$sim(k,l)$  adalah nilai *similarity* antara *item*  $k$  dan *item*  $l$

$m$  adalah jumlah total *user* yang merating *item*  $k$  dan *item*  $l$

$\bar{R}_u$  adalah *rating* yang diberikan oleh *user*  $u$  pada semua *item*

$R_{u,k}$  dan  $R_{u,l}$  adalah *rating* yang diberikan oleh *user*  $u$  kepada *item*  $k$  dan *item*  $l$

Menghitung Prediksi dengan Non Cold Start Problem. Metode *weighted average of deviation* yang didapat dari rata-rata *item* yang telah dirating merupakan metode yang digunakan untuk prediksi *rating* pada *item*  $k$  yang

telah dirating. Rumus berikut ini merupakan perhitungan prediksi rating pada item  $l$  untuk user  $u$ .

$$P_{u,k} = \bar{R}_k + \frac{\sum_{l=1}^n (R_{u,l} - \bar{R}_l) \times \text{sim}(k,l)}{\sum_{l=1}^n |\text{sim}(k,l)|}$$

Keterangan:

$P_{u,k}$  adalah prediksi rating item  $k$  untuk user  $u$

$n$  adalah jumlah rated item user  $u$

$\bar{R}_{u,l}$  adalah rating dari user  $u$  untuk item  $l$

$\bar{R}_k$  dan  $\bar{R}_l$  adalah rating rata-rata untuk item  $k$  dan item  $l$

$\text{Sim}(k,l)$  adalah nilai similarity antara item  $k$  dengan seluruh rated item active user

Cold Start Problem. Metode perhitungan prediksi pada non cold-start problem yaitu weighted average of deviation masih kurang dapat diimplimentasikan pada masalah item baru yang belum dirating karena  $\bar{R}_k$  yang merupakan nilai rata-rata pada item  $k$  akan bernilai nol (karena belum ada yang memberi rating). Oleh karena itu digunakan metode weighted sum untuk menghitung prediksi rating pada kasus item baru. Berikut rumus perhitungannya pada persamaan

$$P_{u,k} = \bar{R}_k + \frac{\sum_{l=1}^n R_{u,l} \times \text{sim}(k,l)}{\sum_{l=1}^n |\text{sim}(k,l)|}$$

Keterangan:

$P_{u,k}$  adalah prediksi rating item  $k$  untuk user  $u$

$n$  adalah jumlah rated item user  $u$

$\bar{R}_{u,l}$  adalah rating diberikan user  $u$  kepada item  $l$

$\text{Sim}(k,l)$  adalah nilai similarity antara item  $k$  dengan seluruh rated item ke- $l$

Akurasi sistem rekomendasi dilihat berdasarkan nilai mean absolute error (MAE)., yaitu rata-rata dari error yang di absolutkan. Dimana error merupakan selisih dari nilai rating sebenarnya dengan nilai rating hasil prediksi. Berikut adalah perhitungan MAE yang ditunjukkan oleh Persamaan.

$$MAE = \frac{\sum_{u=1}^N |P_{u,i} - R_{u,i}|}{N}$$

Dimana  $P_{u,i}$  adalah prediksi rating user  $u$  untuk item  $i$  dan  $R_{u,i}$  adalah nilai rating sebenarnya yang telah diberikan oleh user  $u$  untuk item  $i$ .

## PEMBAHASAN

Dalam penelitian ini dilakukan beberapa pengujian, hasil pengujian yang diperoleh tersebut adalah sebagai berikut:

Pada uji coba 1 dilakukan pengujian dengan data yang digunakan sebanyak 5 user dan 5 buku dengan besarnya rating yang bervariasi. Dari hasil uji coba 1 dapat disimpulkan bahwa hasil prediksi yang dihasilkan oleh sistem cukup akurat dan sama hasilnya dengan prediksi manual yang dihitung oleh Microsoft Excel, ini juga di buktikan oleh kecilnya MAE yang diberikan oleh sistem.

Tabel 1 Tabel rating user terhadap buku

		User				
		1	2	3	4	5
Buku	1	3	0	8	4	5
	2	5	8	7	3	7
	3	7	4	0	7	6
	4	8	7	9	9	8
	5	10	3	8	3	7

Langkah pertama adalah mencari nilai rata-rata rating dari setiap buku.

Tabel 2 Tabel rata-rata rating

		User					rata-rata rating
		1	2	3	4	5	
Buku	1	3	0	8	4	5	4
	2	5	8	7	3	7	6
	3	7	4	0	7	6	4,8
	4	8	7	9	9	8	8,2
	5	10	3	8	3	7	6,2

Langkah kedua adalah mencari nilai rating – (rata-rata rating) lalu dikuadratkan.

Langkah ketiga adalah mencari jumlah dari nilai rating-(rata-rata rating)<sup>2</sup> perbuku dan selanjutnya diakarkan. Terlihat seperti pada Tabel 3.

Tabel 3 Tabel jumlah rating-  
(rata-rata rating)<sup>2</sup> per buku

Sum (rata-rata rating) <sup>2</sup> )	Akar (rating-(rata- rata rating) <sup>2</sup> )	Sum (rata-rata rating) <sup>2</sup> )
34	5,830952	
16	4	
34,8	5,899152	
2,8	1,67332	
38,8	6,228965	

Langkah keempat menghitung *similarity* antar buku dengan persamaan rumus dibawah. Terlihat seperti pada Tabel 5.

$$sim(k,l) = \frac{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)(R_{u,l} - \bar{R}_l)}{\sqrt{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)^2} \sqrt{\sum_{u=1}^m (R_{u,l} - \bar{R}_l)^2}}$$

Keterangan:

*sim(k,l)* adalah nilai *similarity* antara *item k* dan *item l*

*m* adalah jumlah total *user* yang *merating item k* dan *item l*

$\bar{R}_k$  dan  $\bar{R}_l$  adalah *rating* rata-rata pada *item k* dan *item l*

$R_{u,k}$  dan  $R_{u,l}$  adalah *rating* yang diberikan oleh *user u* kepada *item k* dan *item l*

Tabel 4 Tabel similarity

Sim(1,1)	1
sim(1,2)	-0,085749293
sim(2,3)	-0,593305566
sim(3,4)	-0,182349202
sim(4,5)	0,17269415
sim(1,3)	-0,494219459
sim(1,4)	0,819920062
sim(1,5)	0,468051455
sim(2,4)	-0,597614305
sim(2,5)	0,080270162
sim(3,5)	-0,103413708

Setelah diketahui nilai dari *similarity* antar buku, langkah selanjutnya adalah menghitung nilai prediksi buku terhadap user. Hasil dari prediksi manual dapat diimplimentasi kedalam Tabel 5.

Tabel 5 Hasil prediksi manual uji coba 1 menggunakan M. Excel

Hasil Prediksi	User				
	1	2	3	4	5
1	3,86	1,8	6,71	3,41	4,185
2	5,23	7,39	7,3	3,86	6,163
Buku 3	6,03	5,02	1,55	5,49	4,82
4	8,13	6,00	9,88	8,79	8,178
5	7,83	3,43	8,60	4,26	6,852

Tabel 6 Hasil MAE sistem uji coba 1  
MAE

Buku 1	1,07
Buku 2	0,57
Buku 3	1,24
Buku 4	0,48
Buku 5	0,92

ID	BookRating	ISBN	BookRating	User_ID
1	2740230	3.86509	7	
2	2740230	1.81802	8	
3	2740230	6.71447	9	
4	2740230	3.41705	10	
5	2740230	4.18538	11	
6	20199090	5.23843	7	
7	20199090	7.39075	8	
8	20199090	7.3455	9	
9	20199090	3.86154	10	
10	20199090	6.16378	11	
11	20437501	6.03501	7	
12	20437501	5.02754	8	
13	20437501	1.55463	9	
14	20437501	6.55493	10	
15	20437501	4.8279	11	
16	60013117	8.13968	7	
17	60013117	6.0065	8	
18	60013117	9.8837	9	
19	60013117	8.79117	10	
20	60013117	8.17895	11	
21	60168013	7.83867	7	
22	60168013	3.43959	8	
23	60168013	8.6046	9	
24	60168013	4.26506	10	
25	60168013	0.654384	11	

Gambar 1. Hasil prediksi sistem uji coba

Pada uji coba 2 dilakukan uji coba dengan data sebanyak 5 user dan 5 buku dengan besarnya rating yang bervariasi dan terdapat user baru yang belum pernah merating sama sekali. Dari hasil uji coba 2 dapat disimpulkan bahwa hasil prediksi yang dihasilkan oleh sistem cukup akurat dan sama hasilnya dengan prediksi manual yang dihitung oleh Microsoft Excel, ini juga dibuktikan oleh kecilnya MAE yang diberikan oleh sistem. Apabila ada salah satu user yang belum pernah merating sama sekali maka sistem tetap akan memberikan rekomendasinya terhadap user tersebut berdasarkan hasil dari nilai rating yang diberikan oleh user lain yang telah merating.

Tabel 7 Tabel rating user terhadap buku uji coba 2

	User				
	1	2	3	4	5
1	3	0	8	4	0
2	5	8	7	3	0
Buku 3	7	4	0	7	0
4	8	7	9	9	0
5	10	3	8	3	0

Langkah pertama adalah mencari nilai rata-rata rating dari setiap buku. Terlihat seperti pada Tabel 8 dibawah ini

Tabel 8 Tabel rata-rata rating

	User					rata-rata rating
	1	2	3	4	5	
1	3	0	8	4	0	3
2	5	8	7	3	0	4,6
Buku 3	7	4	0	7	0	3,6
4	8	7	9	9	0	6,6
5	10	3	8	3	0	4,8

$$sim(1,2) = \frac{(3-3)(5-4,6) + (0-3)(8-4,6) + (8-3)(7-4,6) + (4-3)(3-4,6) + (0-3)(0-4,6)}{\sqrt{((3-3) + (0-3) + (8-3) + (4-3) + (0-3))^2} \sqrt{((5-4,6) + (8-4,6) + (7-4,6) + (3-4,6) + (0-4,6))^2}}$$

$$sim(1,2) = \frac{(0) + (-10,2) + (12) + (-1,6) + (13,8)}{(\sqrt{(0) + (-3) + (5) + (1) + (-3)}) \sqrt{((0,4) + (3,4) + (2,4) + (-1,6) + (-4,6))}}$$

$$sim(1,2) = \frac{14}{\sqrt{44} \times \sqrt{41,2}}$$

Langkah kedua adalah mencari nilai rating - (rata-rata rating) lalu dikuadratkan.

Langkah ketiga adalah mencari jumlah dari nilai rating-(rata-rata rating)<sup>2</sup> per buku dan selanjutnya diakarkan. Terlihat seperti pada Tabel 9. dibawah ini

Tabel 9 Tabel jumlah rating-(rata-rata rating)<sup>2</sup> per buku

Sum (rating-(rata-rata rating) <sup>2</sup> )	Akar Sum (rating-(rata-rata rating) <sup>2</sup> )
44	6,63325
41,2	6,418723
49,2	7,014271
57,2	7,563068
66,8	8,173127

Langkah keempat menghitung *similarity* antar buku dengan persamaan rumus (2.1). Terlihat seperti pada tabel 4.10

$$sim(k,l) = \frac{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)(R_{u,l} - \bar{R}_l)}{\sqrt{\sum_{u=1}^m (R_{u,k} - \bar{R}_k)^2} \sqrt{\sum_{u=1}^m (R_{u,l} - \bar{R}_l)^2}}$$

Keterangan:

$sim(k,l)$  adalah nilai *similarity* antara *item k* dan *item l*

$m$  adalah jumlah total *user* yang merating *item k* dan *item l*

$\bar{R}_k$  dan  $\bar{R}_l$  adalah *rating* rata-rata pada *item k* dan *item l*

$R_{u,k}$  dan  $R_{u,l}$  adalah *rating* yang diberikan oleh *user u* kepada *item k* dan *item l*

$$sim(1,2) = \frac{14}{6,63325 \times 6,41872}$$

$$sim(1,2) = 0,328816112$$

Tabel 10 Tabel similarity

Sim(1,1)	1
sim(1,2)	0,328816112
sim(2,3)	0,115497394
sim(3,4)	0,531579889
sim(4,5)	0,672987635
	-
sim(1,3)	0,107463533
sim(1,4)	0,657793514
sim(1,5)	0,627139777
sim(2,4)	0,704497544
sim(2,5)	0,545166541
sim(3,5)	0,289559158

Setelah diketahui nilai dari similarity antar buku, langkah selanjutnya adalah menghitung nilai prediksi buku terhadap user. Dapat diimplimentasi kedalam Tabel 11

Tabel 11 Hasil prediksi manual uji coba 2 menggunakan M. Excel

Hasil Prediksi	User				
	1	2	3	4	5
1	4,450	1,974	6,587	3,205	-1,217
2	6,312	5,25	7,2	4,537	-0,325
Buku 3	6,386	3,994	2,789	4,836	-0,659
4	8,559	6,550	8,736	7,308	1,8456
5	7,142	4,339	7,421	4,97	0,1192

ID_BookRating	ISBN	BookRating	User_ID
1	2740230	4.45089	2
2	20199090	6.31265	2
3	20437501	6.386	2
4	60013117	8.5593	2
5	60168013	7.14293	2
6	2740230	1.97445	3
7	2740230	6.58722	4
8	2740230	3.20519	5
9	2740230	-1.21775	6
10	20199090	5.2534	3
11	20199090	7.222	4
12	20199090	4.53727	5
13	20199090	-0.325328	6
14	20437501	3.99455	3
15	20437501	2.78901	4
16	20437501	5.4895	5
17	20437501	-0.659678	6
18	60013117	6.55042	3
19	60013117	8.73623	4
20	60013117	7.30835	5
21	60013117	1.84565	6
22	60168013	4.33974	3
23	60168013	7.42113	4
24	60168013	4.9769	5
25	60168013	0.523982	6

Gambar 2. Hasil prediksi sitem uji coba 2

Tabel 12 Hasil MAE dari uji coba 2

MAE	
Buku 1	1,37
Buku 2	1,23
Buku 3	1,12
Buku 4	0,96
Buku 5	1,46

Pada uji coba 3 dilakukan pengecekan terhadap sistem dengan menggunakan perbandingan terhadap perhitungan manual pada Microsoft Excel. Data yang digunakan sebanyak 5 user dan 6 buku dengan besarnya rating yang bervariasi dan terdapat buku baru yang belum pernah dirating sama sekali. Dari hasil uji coba 3 dapat disimpulkan bahwa apabila terdapat buku baru dan belum pernah dirating sama sekali oleh user maka buku tersebut tidak akan direkomendasikan oleh sistem.



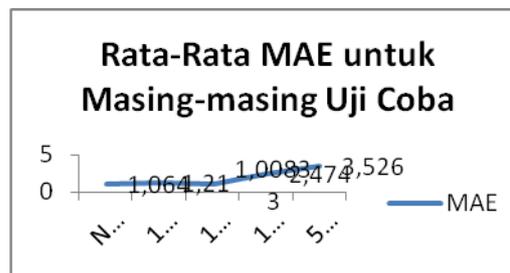
Gambar 3. Buku baru tidak pernah dirating

Pada uji coba 4 dilakukan pengecekan terhadap sistem dengan menggunakan data yang digunakan sebanyak 10 user dan 5 buku dengan besarnya rating yang bervariasi. Dari hasil uji coba 4 dapat disimpulkan bahwa semakin banyak user yang menggunakan sistem maka hasil dari prediksi yang dihasilkan kurang akurat.



Gambar 4 Hasil prediksi kurang akurat

Pada uji coba 5 dilakukan pengecekan terhadap sistem dengan menggunakan data yang digunakan sebanyak 5 user dan 10 buku dengan besarnya rating yang bervariasi. Pada Tabel 4.27 digambarkan besarnya rating yang diberikan oleh user terhadap 5 buah buku. Dari hasil uji coba 5 dapat disimpulkan bahwa semakin banyak data, dalam hal ini adalah buku yang digunakan sebanyak 10 dan jika dibandingkan dengan percobaan 1 yaitu 5 buku maka hasil rekomendasi yang dihasilkan kurang baik.



Gambar 5 Hasil MAE uji coba 1,2,3,4 dan 5

## KESIMPULAN

Metode *collaborative filtering* dapat diimplementasikan dalam pembuatan sistem rekomendasi buku dengan melihat kedekatan buku berdasarkan nilai rating. Metode ini lemah ketika diimplementasikan pada buku baru yang belum pernah dirating sama sekali. Hasil prediksi *rating* setiap buku untuk masing-masing user dengan menggunakan metode *collaborative filtering* kurang baik. Hal ini ditunjukkan berdasarkan rata-rata nilai MAE (*Mean Absolute Error*) buku pada uji coba 1 yakni 1,064 lebih kecil dari pada uji coba 2 yakni 1,21, uji coba 4 yakni 2,474 dan ujicoba 5 yakni 3,526. Hal ini menunjukkan bahwa semakin banyak jumlah data yang digunakan dan jika terdapat user yang belum pernah merating, maka sistem yang dihasilkan relatif tidak akurat dan menghasilkan rekomendasi yang buruk. Oleh karena itu pada penelitian selanjutnya, metode *collaborative filtering* diharapkan agar diimplementasikan pada data yang memiliki *item* yang banyak dirating oleh user. Apabila terdapat data yang banyak dan memiliki *item* baru yang sedikit dirating oleh user, maka diharapkan menggunakan metode yang lebih baik dari *collaborative filtering*, misalnya adalah ICHM (*Item-Based Clustering Hybrid Method*). ICHM (*Item-Based Clustering Hybrid Method*) adalah salah satu metode yang menggunakan pendekatan *hybrid* atau menggabungkan kedua pendekatan yaitu *Content Based Filtering* dan *Collaborative Filtering*.

## DAFTAR PUSTAKA

Djamal, A Rhamadanus. Maharani, Warih dan Kurniati, Angelina

- Prima (2010). *Analisis dan Implementasi Metode Item-Based Clustering Hybrid Pada Recomennder Sytem.*
- Lam, S. And Riedl, J. (2004). Shilling recommender systems for fun and profit. In *Proceedings of the 13th International WWW Conference.* New York..
- Li, Qing and Kim, Byeong Man 2002. *An Approach for Combining Content-based and Collaborative Filters.* Departement of Computer Sciences, Kumoh National Institute of Technology.
- Sarwar, Badrul et al. 2001. *Item-based Collaborative Filtering Recommender System Algorithm.* GroupLens Research Group/Army HPC Research Center, Department of Computer Science and Engineering, University of Minnesota. Mienneapolis.
- Uyun, S. Fahrurrozi, I. dan Mulyanto, A. 2011. *Item Collaborative Filtering untuk Rekomendasi Pembelian Buku secara Online.* JUSI, Vol. 1, No. 1 ISSN 2087-8737