

Peningkatan Efektivitas Metode *User-item based Collaborative Filtering* pada Sistem Rekomendasi Wisata Kuliner Kota Solo

Sayekti Hadi Ati
Jurusan Informatika
Universitas Sebelas Maret
Jalan Ir.Sutami 36 A
Surakarta 57126
say2_five@yahoo.com

Ristu Saptono
Jurusan Informatika
Universitas Sebelas Maret
Jalan Ir.Sutami 36 A
Surakarta 57126
r_saptono@uns.ac.id

Umi Salamah
Jurusan Informatika
Universitas Sebelas Maret
Jalan Ir.Sutami 36 A
Surakarta 57126
u_salamah@yahoo.com

ABSTRAK

Salah satu teknik rekomendasi yang telah sukses dan banyak digunakan yaitu *collaborative filtering*. *Collaborative filtering* dapat dibagi menjadi dua metode utama yaitu *user based* dan *item based*. Pada umumnya kedua metode tersebut belum memiliki fitur untuk mengakomodasi item baru. Oleh karena itu digunakan metode alternatif yang disebut *user-item based collaborative filtering*. Metode ini memberikan rekomendasi berdasarkan urutan teratas nilai prediksi *rating* yang dihasilkan oleh sistem. Metode ini menggunakan algoritma *missing-value* pada perhitungan prediksi *rating* dan tidak memerlukan proses pembentukan *neighborhood*. Kelemahan dari metode *user-item based collaborative filtering* adalah ketika pengujian dilakukan dengan pengukuran *error* menggunakan *normalized mean absolute error* (NMAE), hasil yang diperoleh NMAE cukup tinggi. Dengan melakukan perbaikan pada rumus prediksi *missing value algorithm* (MVA) menjadi *adjusted MVA*, metode *user-item based collaborative filtering*, dalam pengujian terbukti dapat menurunkan NMAE dan mengurangi jumlah prediksi di luar *range* nilai *rating*. Sebagai studi kasus metode *user-item based* ini diterapkan pada sistem rekomendasi wisata kuliner Kota Solo.

Kata kunci

Algoritma *missing value*, sistem rekomendasi, *user-item based collaborative filtering*, prediksi *rating*.

1. PENDAHULUAN

Salah satu jenis pariwisata yang terkenal dari Kota Solo adalah wisata kuliner. Wisatawan akan dibingungkan oleh banyaknya pilihan kuliner yang ditawarkan. Terbatasnya waktu berkunjung membuat wisatawan tidak mungkin mencoba semua tempat wisata kuliner satu per satu sehingga memerlukan rekomendasi yang berasal dari orang lain maupun dari sistem. Sistem rekomendasi merupakan salah satu sistem yang dapat memberikan solusi permasalahan tersebut.

Untuk mengatasi permasalahan tersebut, penulis merancang dan membuat sistem rekomendasi wisata kuliner Kota Solo. Dalam pembuatan sistem rekomendasi ini, penulis menggunakan salah satu teknik rekomendasi yang telah sukses dan banyak digunakan yaitu *collaborative filtering*. *Collaborative filtering* merupakan teknik yang menggunakan preferensi diketahui dari sekelompok pengguna untuk memprediksi preferensi yang tidak diketahui dari pengguna baru; rekomendasi untuk pengguna baru tersebut berdasar pada prediksi ini [5].

Collaborative filtering dapat dibagi menjadi dua metode utama yaitu *memory based (user based)* dan *model based (item based)*. Metode *memory based* atau *user based* memanfaatkan riwayat tentang pilihan pengguna. Untuk kasus wisata kuliner, jika ada pengguna A yang menyukai kuliner K, sedangkan profil

pengguna A mirip dengan profil pengguna B dan pengguna B belum pernah memberi penilaian pada kuliner K, maka kuliner K akan direkomendasikan kepada pengguna B. Pendekatan *user based* ini mempunyai kelebihan yaitu mampu menghasilkan rekomendasi berkualitas tinggi, namun mempunyai kelemahan yaitu kompleksitas perhitungan bertambah seiring dengan bertambahnya pengguna dan item [6].

Untuk metode *model based* atau *item based* menggunakan kemiripan riwayat nilai *rating* item. Jadi, jika terdapat pengguna A menyukai kuliner K, sedangkan nilai *rating* kuliner K mirip dengan kuliner L dan kuliner L belum pernah di-*rating* oleh A, maka kuliner L dapat direkomendasikan kepada A. Metode ini mengasumsikan kesamaan antara dua item berpusat pada satu nilai sehingga mudah menemukan *neighbor* pada sistem *offline*. Karena inilah metode *model based* atau *item based* ini mampu menghasilkan rekomendasi dengan cepat namun sayangnya kualitas rekomendasi yang dihasilkan rendah [6].

Pada umumnya *collaborative filtering* belum memiliki fitur untuk mengakomodasi item baru. Hal ini dapat dipahami karena sebuah item baru yang belum pernah mendapatkan *rating* sebelumnya tidak dapat dihitung *similarity* atau kesamaannya dengan item lainnya. Oleh karena itu digunakan metode alternatif yang dapat menangani kelemahan dari dua metode utama di atas yang disebut dengan metode *user-item based collaborative filtering*. Metode ini dapat menghasilkan rekomendasi tanpa mencari formasi ketetanggaan karena menggunakan algoritma *missing value* untuk menghitung prediksi *rating* dan mampu merekomendasikan item baru pada sistem [5].

Dari paparan di atas maka dipilih metode *user-item based collaborative filtering* dalam pembuatan sistem rekomendasi ini. Dengan metode tersebut sistem ini diharapkan dapat memberikan rekomendasi khususnya merekomendasikan item baru yang belum pernah di-*rating* kepada pengguna. Selain itu sistem ini juga dapat memberikan rekomendasi kepada pengguna baru dengan atau tanpa melalui *training set*. Pemberian *rating* oleh pengguna menggunakan *continuous rating bar*. Pemilihan menggunakan *continuous rating bar* karena memiliki kelebihan antara lain: *continuous rating* menghindari efek diskritisasi dalam perhitungan matriks, *continuous rating* menangkap rasa dengan rincian yang lebih halus, dan pengguna melaporkan bahwa *continuous rating bar* mudah untuk digunakan [1].

Setelah pembangunan sistem selesai dan diuji coba dengan data *real*, keefektifannya diuji dengan pengukuran *error* menggunakan *Normalized Mean Absolute Error* (NMAE). Dari pengujian tersebut diperoleh hasil bahwa nilai NMAE cukup tinggi karena selisih antara prediksi *rating* dengan *rating*

sebenarnya cukup jauh. Oleh karena itu perlu adanya perbaikan pada rumus prediksi metode *user-item based collaborative filtering* agar nilai NMAE dapat diturunkan dan mengurangi jumlah prediksi *rating* di luar *range*.

2. LANDASAN TEORI

2.1 Sistem Rekomendasi Berbasis Collaborative Filtering

Sistem rekomendasi merupakan model aplikasi dari hasil observasi terhadap keadaan dan keinginan pengguna. Sistem rekomendasi memerlukan model rekomendasi yang tepat agar apa yang direkomendasikan sesuai dengan keinginan pengguna, serta mempermudah pengguna mengambil keputusan yang tepat dalam menentukan item yang akan dipilih [3].

Sistem rekomendasi memberikan saran menurut selera atau ukuran tertentu tentang item yang mungkin menarik minat pengguna dengan cara mencocokkan pilihan dengan profil pengguna atau grup pengguna. Sistem ini harus cerdas dan bijaksana memahami apa yang pengguna inginkan. Dengan kata lain, sistem cukup bijak untuk mempelajari penggunaannya, membuat profil mereka, dan memberikan rekomendasi atau saran [2].

2.2 Collaborative Filtering

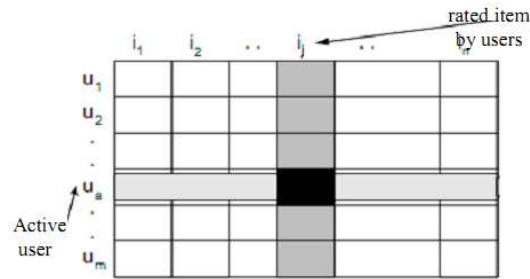
Collaborative filtering (CF) adalah teknik bagaimana membuat prediksi otomatis atau rekomendasi tentang minat seorang pengguna dengan pengumpulan informasi cita rasa dari banyak pengguna. Prinsip kerja dari CF adalah memberikan prediksi atau rekomendasi item berdasarkan pada opini pengguna-pengguna yang memiliki kemiripan [6].

Tujuan dari CF adalah merekomendasikan item baru atau memprediksi evaluasi item kepada pengguna tertentu berdasarkan pada ketertarikan sebelumnya dari pengguna tersebut dan opini pengguna-pengguna lain yang mempunyai ketertarikan yang mirip. Di dalam CF terdapat daftar *m* pengguna $U = \{u_1, u_2, \dots, u_m\}$ dan daftar *n* item $I = \{i_1, i_2, \dots, i_n\}$. Setiap pengguna u_i mempunyai daftar item I_{u_i} yang menggambarkan opini dari pengguna tersebut. Opini diberikan secara eksplisit oleh pengguna berupa suatu nilai *rating*, biasanya dalam skala kuantitatif tertentu atau bisa juga secara implisit yang diperoleh dari arsip transaksi, dengan menganalisis *logs*, *navigation history*, atau dengan cara yang lain [6].

Ada dua hal utama yang dilakukan sistem rekomendasi dengan *collaborative filtering*, yaitu prediksi dan rekomendasi. Prediksi merupakan penilaian yang diberikan sistem pada suatu item sebagai perkiraan terhadap penilaian yang diberikan oleh pengguna. Rekomendasi merupakan daftar item dengan nilai prediksi tertinggi. Item-item yang direkomendasikan belum pernah di *rating* sebelumnya oleh pengguna tersebut. Antarmuka ini pada sistem rekomendasi dengan *collaborative filtering* disebut *Top-N recommendation* [4].

2.3 User-item based Collaborative Filtering

Metode *user-item based collaborative filtering* tidak seperti metode lainnya. Metode ini melewati proses *neighborhood formation*. Pada metode ini dihasilkan sebuah matriks oleh pengguna *u* yang memberi *rating* pada item *i* [5].



Gambar 1. Matriks *rating user-item* (Saptono, 2010)

Matrik di atas dianggap sebagai rancangan acak tak lengkap kemudian nilai yang hilang dapat dihitung langsung dengan menggunakan algoritma *missing value (MVA)*. Misalkan matriks *rating user-item* memiliki *m* pengguna dengan *n* item, prediksi *rating* untuk item *j* oleh pengguna *i* $P(i,j)$ dapat ditulis sebagai berikut

$$P(i, j) = \frac{n * R'(i, *) + m * R'(*, j) - R'(*, *)}{(m - 1)(n - 1)}$$

Dimana $R'(i, *)$ menyatakan jumlah *rating* pada baris pengguna *i*, $R'(*, j)$ menyatakan jumlah *rating* pada kolom item *j* dan $R'(*, *)$ menyatakan jumlah semua *rating* pada matriks.

Perhitungan dengan MVA menyebabkan banyaknya nilai prediksi yang berada di luar jangkauan *rating* yang diberikan oleh pengguna. Oleh karena itu, untuk memperbaiki nilai prediksi digunakan MVA yang disesuaikan (*adjusted MVA*).

Sedikit berbeda perhitungan prediksi *rating* dengan MVA yang memperhitungkan keseluruhan *rating* dalam *user-item rating matrix*. Pada perhitungan prediksi menggunakan *adjusted MVA*

hanya memperhitungkan sel yang memiliki *rating*. Perhitungan dilakukan dengan mengalikan jumlah sel yang berisi *rating* dengan rata-rata nilai *rating* pada *user-item rating matrix*.

$$P(i, j) = \left[\frac{n * R'(i, *) + m * R'(*, j) - \left(\frac{R'(*, *)}{mn} \right) (mn - sparse)}{(m - 1)(n - 1)} \right]$$

Di mana *n* menyatakan jumlah item, *m* menyatakan jumlah pengguna, *mn* menyatakan perkalian jumlah item dan jumlah pengguna, (*mn-sparse*) menunjukkan jumlah sel *rating* yang terisi (*rating* tidak sama dengan nol), $R'(i, *)$ menyatakan jumlah *rating* pada pengguna *i*, $R'(*, j)$ menyatakan jumlah *rating* pada item *j* dan $R'(*, *)$ menyatakan jumlah semua *rating* pada matriks.

3. PEMBAHASAN

3.1 Evaluasi Sistem

Evaluasi atau pengukuran kualitas dari sistem rekomendasi dilakukan dengan menggunakan *Mean Absolute Error (MAE)*. MAE adalah suatu ukuran penyimpangan rekomendasi dari nilai benar *user-specified* [6]. Jika p_{ui} adalah nilai prediksi *rating* yang diberikan pengguna *u* pada item *i*, dan r_{ui} adalah nilai *rating* yang sebenarnya maka persamaan MAE dirumuskan sebagai berikut:

$$MAE = \frac{1}{c} \sum_{i=1}^c |r_{ui} - p_{ui}|$$

Dimana c adalah jumlah item yang direkomendasikan dan telah di-rating oleh pengguna u . Semakin kecil nilai MAE, semakin akurat sistem dalam memberikan rekomendasi.

Salah satu alternatif pengukuran *error* dengan menggunakan *normalized mean absolute error* (NMAE) [1]. Dengan skala *rating* yang diberikan pada interval [1,10] maka NMAE dapat dihitung dengan persamaan:

$$NMAE = \frac{MAE}{r_{max} - r_{min}}$$

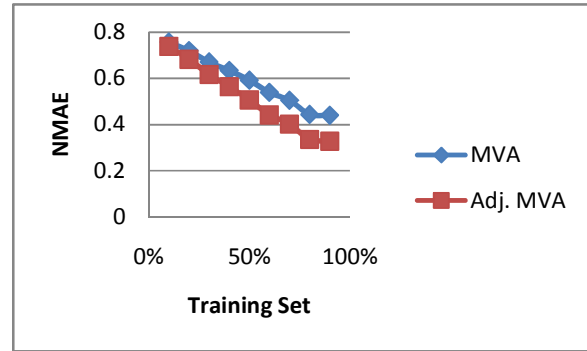
Dengan r_{max} adalah nilai *rating* maksimal yang diberikan oleh pengguna, sedangkan r_{min} adalah nilai *rating* minimal pada skala [1,10].

3.2 Hasil Pengujian

Pada bagian ini disajikan hasil pengujian metode *user-item based collaborative filtering* pada sistem rekomendasi wisata kuliner Kota Solo menggunakan *missing value algorithm* (MVA) dan *adjusted MVA*. Pengujian dilakukan pada data sampel *user-item rating matrix* yang terdiri dari 26 pengguna dan 14 item kuliner dengan jumlah sel *rating* yang terisi sebanyak 82 *rating* dengan tingkat *sparsity* 0.774. Data yang digunakan untuk pengujian hanya data yang nilai *rating* nya terisi atau nilainya tidak sama dengan nol. Pengujian dilakukan sebanyak 30 kali dengan pengosongan *rating* secara acak mulai dari 10%, 20%, sampai dengan 90%. Setelah dilakukan 30 kali pengujian, data yang terkumpul dihitung rata-rata NMAE dan jumlah prediksi *rating* yang nilainya di luar *range* untuk setiap tingkat pengosongan *rating*. Tabel dan grafik berikut ini menunjukkan perbandingan rata-rata nilai NMAE dan rata-rata jumlah prediksi di luar *range*.

Tabel 1. Perbandingan Rata-rata NMAE Rumus Prediksi MVA vs Rumus Prediksi *Adjusted MVA*

Training Set	Skala Sparsity	MVA	Adjusted MVA
10%	0.98	0.757	0.738
20%	0.96	0.72	0.683
30%	0.93	0.671	0.617
40%	0.91	0.634	0.564
50%	0.89	0.592	0.507
60%	0.87	0.539	0.442
70%	0.84	0.505	0.402
80%	0.82	0.444	0.336
90%	0.80	0.44	0.328



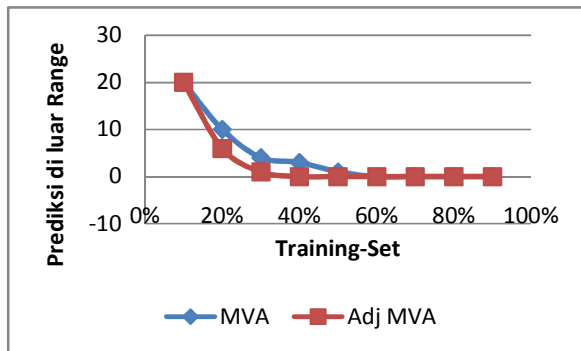
Gambar 2. Perbandingan NMAE antara rumus MVA dan rumus *Adjusted MVA*

Dari tabel dan grafik di atas menunjukkan bahwa metode *user-item based collaborative filtering* menggunakan rumus prediksi *adjusted MVA* menghasilkan kualitas prediksi yang lebih baik dibandingkan menggunakan rumus prediksi MVA. Hal ini bisa dilihat dari nilai NMAE yang dihasilkan oleh rumus prediksi *adjusted MVA* lebih kecil daripada nilai NMAE yang dihasilkan oleh rumus prediksi MVA. Semakin kecil nilai NMAE menunjukkan semakin baik kualitas prediksi yang dihasilkan oleh sistem. Semakin besar tingkat pengosongan *rating*, nilai NMAE yang dihasilkan semakin besar.

Kualitas rekomendasi akan berkurang jika nilai prediksi yang dihasilkan berada jauh dari nilai sebenarnya. Oleh karena itu perlu dilakukan perhitungan banyaknya sel yang nilai prediksinya berada di luar domain atau *range* prediksi yaitu lebih besar dari 5 (lima) dan kurang dari atau sama dengan 0 (nol).

Tabel 2. Perbandingan Rata-rata Jumlah Prediksi di luar *Range* dari Rumus MVA vs Rumus *Adjusted MVA*

Training Set	Skala Sparsity	MVA	Adjusted MVA
10%	0.98	20	20
20%	0.96	10	6
30%	0.93	4	1
40%	0.91	3	0
50%	0.89	1	0
60%	0.87	0	0
70%	0.84	0	0
80%	0.82	0	0
90%	0.80	0	0



Gambar 3. Perbandingan Prediksi di luar Range antara Rumus MVA dan Rumus Adjusted MVA

Untuk jumlah prediksi di luar range, rumus adjusted MVA menghasilkan jumlah yang lebih sedikit dibandingkan dengan rumus MVA. Pada tingkat pengosongan rating sebanyak 90%, kedua rumus menghasilkan rata-rata jumlah prediksi di luar range yang sama besar.

3.3 Analisis Hasil Pengujian

Pengujian metode *user-item based collaborative filtering* dengan dua rumus prediksi yang berbeda pada sampel sebanyak 82 rating yang dilakukan sebanyak 30 kali dengan pengosongan acak dari tingkat 10% sampai dengan 90% diperoleh hasil bahwa metode *user-item based collaborative filtering* yang menggunakan rumus prediksi adjusted MVA menghasilkan kualitas prediksi yang lebih baik. Prediksi rating yang baik menghasilkan kualitas rekomendasi yang baik pula.

Rumus prediksi adjusted MVA mampu menurunkan rata-rata nilai NMAE sebesar 25 % dari rata-rata nilai NMAE yang dihasilkan oleh rumus prediksi MVA pada tingkat pengosongan 10%. Sedangkan pada tingkat pengosongan 90%, rumus prediksi adjusted MVA hanya mampu menurunkan rata-rata nilai NMAE sebesar 3% dari rata-rata nilai NMAE yang dihasilkan rumus prediksi MVA.

Semakin besar tingkat pengosongan rating, semakin besar pula nilai NMAE. Hal ini menunjukkan bahwa metode *user-item based collaborative filtering* mengalami penurunan kualitas prediksi ketika menangani data dengan *sparsity* yang besar. Hal ini bisa dilihat dari rata-rata jumlah prediksi di luar range yang dihasilkan oleh kedua rumus. Pada tingkat pengosongan 50% hingga 80% rumus adjusted MVA mampu mengurangi rata-rata jumlah prediksi di luar range yang dihasilkan oleh rumus prediksi MVA. Namun pada tingkat pengosongan 90%, rata-rata jumlah prediksi di luar range jumlahnya sama besar untuk kedua rumus prediksi tersebut.

Berdasarkan analisis di atas, metode *user-item based collaborative filtering* dengan rumus prediksi adjusted MVA

memberikan hasil kualitas rekomendasi yang lebih baik ditinjau dari rata-rata nilai NMAE dan rata-rata jumlah prediksi di luar range.

4. PENUTUP

Berdasarkan hasil pengujian metode *user-item based collaborative filtering* dengan menggunakan rumus prediksi MVA (*missing value algorithm*) dan rumus prediksi MVA yang sudah dimodifikasi menjadi *adjusted MVA* diperoleh kesimpulan bahwa metode *user-item based collaborative filtering* dengan rumus prediksi adjusted MVA memberikan kualitas rekomendasi yang lebih baik. Hal tersebut dibuktikan pada hasil pengujian NMAE dan hasil prediksi di luar range yang dihasilkan oleh kedua rumus prediksi.

Dari segi implementasi, sistem rekomendasi wisata kuliner Kota Solo yang menggunakan adjusted MVA dalam perhitungan prediksinya menghasilkan urutan rekomendasi yang sama dengan yang dihasilkan oleh sistem rekomendasi wisata kuliner Kota Solo menggunakan rumus prediksi MVA. Jadi dapat diperoleh kesimpulan juga bahwa adjusted MVA meningkatkan kualitas rekomendasi dengan cara menghasilkan nilai prediksi yang lebih baik dari yang dihasilkan MVA tanpa mengubah urutan rekomendasi yang diberikan kepada pengguna.

5. DAFTAR PUSTAKA

- [1] Goldberg, K., Roeder, T., Gupta D., and Perkins, C. 2001. *Eigentaste : A Constant Time Collaborative Filtering Algorithms*, Information Retrieval Journal.
- [2] Kangas, S. 2002. "Collaborative Filtering and Recommendation Systems". LOUHI Project Research Report, VTT Information Technology, Espoo, Finland.
- [3] McGinty, L., Smyth, B. Adaptive. 2006. *selection: analysis of critiquing and preference based feed back in conversation on recommender systems*. International Journal of Electronic Commerce Vol.11 Number 2 / Winter 2006-7 Page 35-57.
- [4] Saptono, R. 2006. *Recommender System untuk Pencarian Buku dengan User Item Based Collaborative Filtering*. Tesis. Bandung : ITB.
- [5] Saptono, R. 2010. *User-item based Collaborative Filtering for Improved Recommendation*. Proceeding ICOSic 2010 UNS Solo Indonesia.
- [6] Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl. 2001. *Item-based Collaborative Filtering Recommendation Algorithms*. 10th International World Wide Web Conference (WWW10 Hongkong, 1-5 Mei 2001), ACM Press, New York, 285-295