

REKOMENDASI PAKET PEMBELIAN BARANG PADA TOKO ONLINE DENGAN COLLABORATIVE FILTERING

Devi Dwi Purwanto

Sistem Informasi, Sekolah Tinggi Teknik Surabaya
devi@stts.edu

ABSTRACT

Currently there are many case found that sale of goods is not only do in traditional method, where buyers and sellers meet, but also can be done by online. To increase sales, seller offers products that are sold with various promos, including the sale of a package with cheaper price when compared to sales of the products separately. However, the promo that seller offers, must known by the seller which product is suitable to made its bundle. Where the packages offered for each customer can be different. The differences can seen by the behavior of customers who give ratings on certain products, and what products are usually purchased together with its products. Collaborative filtering is used to provide a recommendation package purchase of goods for each customer. This is done in addition to increasing sales turnover and provide recommendations on the customer so that the customer is satisfied because the online stores know the needs of the customer.

Keywords : Collaborative Filtering, Link Analysis, Recommendation

ABSTRAK

Saat ini banyak ditemui penjualan barang tidak hanya bersifat tradisional, dimana pembeli dengan penjual bertemu, namun dapat pula penjualan dilakukan secara online. Untuk meningkatkan penjualan, penjual menawarkan produk-produk yang dijual dengan berbagai macam promo, diantaranya penjualan paket dengan harga yang lebih murah bila dibandingkan dengan penjualan produk secara terpisah. Namun demikian penawaran promo tersebut harus diketahui oleh penjual, produk apakah yang cocok dibuat paket. Dimana paket yang ditawarkan oleh tiap pelanggan dapat berbeda. Perbedaan tersebut dapat melihat perilaku pelanggan yang memberi rating pada produk tertentu, serta produk apa yang biasanya dibeli bebarengan dengan produk tersebut. Collaborative filtering digunakan untuk memberikan rekomendasi paket pembelian barang untuk masing-masing pelanggan. Hal ini dilakukan selain untuk meningkatkan omzet penjualan dan memberikan rekomendasi pada pelanggan supaya pelanggan merasa puas karena took online tersebut mengetahui kebutuhan pelanggan.

Kata kunci : Collaborative Filtering, Link Analysis, Rekomendasi

PENDAHULUAN

Seiring dengan perkembangan dunia e-commerce, saat ini penjualan barang banyak dilakukan pada dunia maya dengan menggunakan website sebagai medianya. Untuk meningkatkan penjualan biasanya diberikan penawaran produk lain. Namun rekomendasi barang tersebut kadang tidak sesuai dengan kebutuhan pengguna, sehingga pengguna sering mengabaikan penawaran tersebut. Pada bagian ini peneliti akan mencoba untuk memberikan rekomendasi yang sesuai dengan kebutuhan dan profil user. Profil user sendiri dapat dilihat berdasarkan barang yang diberi rating oleh pengguna. Dengan demikian, pihak penjual dapat melakukan penawaran barang yang dibeli secara bersamaan dengan memberikan potongan harga.

Berangkat dari tujuan yang hendak dicapai tersebut, peneliti akan mengambil profil masing-masing pengguna dengan menggunakan collaborative filtering, sedangkan untuk menentukan paket yang dibeli bersamaan didapatkan dari history para pembeli sebelumnya yang telah melakukan transaksi. Rule pembelian barang yang dibeli secara bersamaan didapatkan dengan menggunakan link analysis.

TINJAUAN PUSTAKA

Collaborative Filtering

Collaborative Filtering merupakan sub bagian dari machine learning yang ditujukan untuk melakukan prediksi keinginan pengguna berdasarkan kegiatan yang dilakukan sekelompok pengguna seperti pemberian rating terhadap suatu barang. Metode collaborative filtering pada prinsipnya adalah proses menyaring data berdasarkan kemiripan informasi, karakteristik, atau profil dari para pengguna.

Teknik pada collaborative filtering antara lain collaborative, content-based, demographic, utility-base, dan knowledge-based. User based collaborative bekerja dengan cara melihat rating yang diberikan oleh pengguna terhadap barang tertentu, dan mengidentifikasi pengguna mana yang kemiripan dengan pengguna tersebut. Content-based bekerja dengan cara melihat fitur atau deskripsi dari barang tanpa mempedulikan user lainnya. Demographic bekerja dengan cara membaca karakteristik pengguna yang berupa data dirinya, kemudian mengidentifikasi karakteristik yang hampir sama. Utility-based bekerja dengan cara membaca fitur dari barang dan melihat pengetahuan tentang bagaimana barang memenuhi kebutuhan pengguna. Knowledge-based bekerja dengan cara menggunakan pendekatan berbasis pengetahuan seperti pengetahuan tentang bagaimana barang tertentu memenuhi kebutuhan pengguna.

Pada user based collaborative memiliki tujuan untuk mencari pengguna mana yang bisa dipercaya untuk dipakai dalam acuan rekomendasi dan pengguna mana yang harus diabaikan dalam pemberian rekomendasi. Untuk mendapatkan acuan rekomendasi tersebut dilakukan dengan cara menghitung korelasi Pearson dimana rumusnya adalah

$$PC(u,v) = \frac{\sum_{i \in \mathcal{I}_{uv}} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in \mathcal{I}_{uv}} (r_{ui} - \bar{r}_u)^2 \sum_{i \in \mathcal{I}_{uv}} (r_{vi} - \bar{r}_v)^2}}$$

LINK ANALYSIS

Link analysis adalah salah satu teknik untuk menganalisa hubungan, pada kasus ini adalah hubungan barang yang dibeli oleh pelanggan pada waktu yang bersamaan. Algoritma yang akan digunakan dalam menentukan rule asosiasi pada penelitian ini adalah algoritma apriori. Dimana algoritma ini dikembangkan oleh IBM untuk mendapatkan rule pada data transaksi yang berjumlah besar. Algoritma apriori dapat dilihat pada algoritma 1.

Algoritma 1: Apriori

```
L1 := {frequent 1-itemsets};  
for ( k := 2; Lk-1 0; k++) do {  
    Ck = apriori-gen(Lk-1);  
    for semua transaksi pada dataset do {  
        for semua kandidat c Ck yang terdapat pada  
        transaksi (t) do  
            c.count++  
        }  
    Lk = { c Ck | c.count >= min-support }  
}  
Hasil :=  $\bigcup_k L_k$ 
```

Pada proses pembuatan apriori-gen dilakukan proses join dan prune untuk mendapatkan kandidat large itemset. Prune dilakukan dengan maksud menghilangkan kandidat dimana penyusunnya tidak termasuk dalam large itemset-1. Dengan dilakukan prune diharapkan penghitungan frekuensi hanya dilakukan pada kandidat yang memenuhi persyaratan sehingga dapat lebih efisiensi waktu.

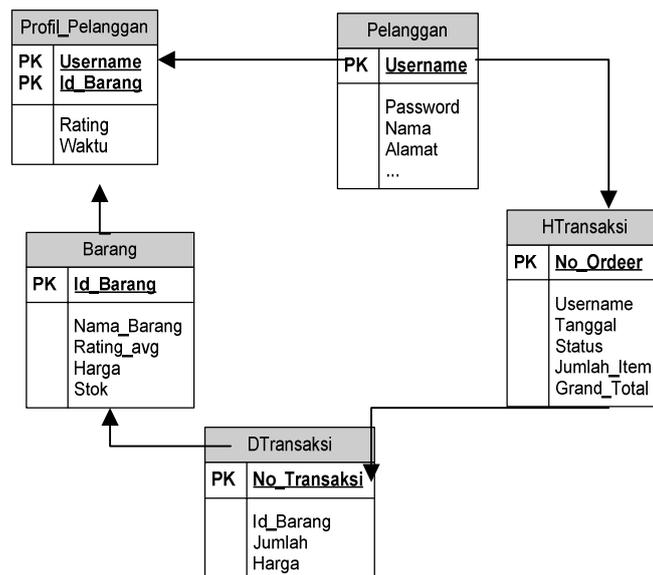
Untuk mendapatkan large itemset perlu ditentukan dahulu nilai untuk minimum support, hasil join dan prune yang memenuhi nilai minimum support akan menjadi large itemset-i. Sedangkan inimum confidence digunakan untuk mendapatkan strong association rulenya. Minimum support digunakan untuk menentukan rule mana yang akan menjadi large itemsetnya. Sedangkan minimum confidence digunakan untuk mendapatkan strong association rule. Rumus dari minimum support dan minimum confidence adalah sebagai berikut:

$$\text{min support}(x) = \frac{\text{jum}(x)}{\text{total transaksi}}$$

$$\text{min conf}(a \rightarrow b, c) = \frac{\text{sup}(a, b, c)}{\text{sup}(a)}$$

DATA

Data yang disediakan pada penelitian ini adalah data transaksi sebanyak 500 transaksi yang dilakukan oleh pelanggan yang berbeda. Untuk mendapatkan profil dari pelanggan, dilakukan pencatatan rating yang dilakukan diberikan oleh pelanggan tertentu pada barang tertentu, sedangkan untuk pencatatan yang lain pada penelitian ini akan diabaikan. Relasi data yang digunakan pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Relasi Data yang Digunakan

Rating yang diberikan oleh pelanggan untuk masing-masing barang akan dicatat pada table profil_pelanggan. Sedangkan field rating pada table barang adalah rata-rata rating barang tersebut yang telah diberikan pada pelanggan. Transaksi penjualan dicatat pada table htransaksi dan dtransaksi. Nantinya table htransaksi dan dtransaksi inilah yang digunakan pada proses pencarian association rule, sedangkan untuk perhitungan collaborative filtering akan menggunakan table profil_pelanggan.

PEMBAHASAN

Profil dari tiap pelanggan tersebut akan dilakukan penghitungan rating terlebih dahulu. Rating diberikan oleh pelanggan untuk barang tertentu. Pelanggan hanya dapat memberikan 1 nilai rating pada sebuah barang. Pada tabel 1 diberikan contoh pemberian rating pada barang Pampers. Dimana rating yang diberikan mempunyai range 1-5.

Tabel 1. Contoh Rating pada Pampers

User	Rating
A	3
B	4
C	5
D	3
E	5

Dari rating yang diberikan tersebut, sistem akan secara otomatis melakukan perhitungan dan mencari rata-rata rating untuk barang pampers. Hasil kalkulasi ditemukan bahwa rating untuk pampers adalah 4 (didapat dari $(3+4+5+3+5)/5$). Hasil tersebut yang akan disimpan pada tabel barang field rating_avg.

User-based Collaborative Filtering bekerja dengan memprediksi rating dari sebuah pelanggan yang memiliki kemiripan pola rating dengan pelanggan tertentu. Pada tabel 2 diberikan contoh rating yang diberikan oleh masing-masing pelanggan terhadap barang yang dijual. Dimana barang yang tidak diberi rating oleh pelanggan akan dianggap memiliki rating 0.

Tabel 2. Contoh Rating User-Barang

User	Brg 1	Brg 2	Brg 3	Brg 4	Brg 5
A	5	1		2	2
B	1	5	2	5	5
C	2	?	3	5	4
D	4	3	5	3	-

Perhitungan kemiripan akan dihitung dengan menggunakan tingkat korelasi Pearson. Dari contoh rating pada tabel 2 tersebut didapatkan masing-masing korelasi antara pelanggan yang satu dengan pelanggan yang lain. Contoh perhitungan korelasi tersebut dapat dilihat pada tabel 3.

Tabel 3. Contoh Hasil Korelasi

User	A	B	C	D
A	1	-0.94	-0.84	0.66
B	-0.94	1	0.92	-0.79
C	-0.84	0.92	1	-0.66
D	0.66	-0.79	-0.66	1

Dari contoh hasil perhitungan korelasi tersebut didapatkan bahwa pelanggan yang memiliki profil yang hampir sama adalah pelanggan B dan pelanggan C, sedangkan pelanggan yang paling tidak sama profilnya adalah pelanggan A dengan pelanggan B. Dimana nilai korelasi semakin mendekati nilai 1 berarti similaritynya semakin besar. Korelasi pelanggan tersebut tidak

diambil untuk satu pelanggan saja, namun bisa lebih dari satu. Hal ini terjadi apabila tingkat korelasi masih berada dalam range threshold yang ditentukan.

Dari korelasi tersebut langkah selanjutnya adalah mendapatkan data pelanggan tersebut pernah membeli barang apa saja yang diambil dari table htransaksi dan dtransaksi. Data transaksi tersebut yang akan diambil untuk diproses dengan algoritma apriori. Contoh data barang yang dibeli secara bersamaan untuk beberapa transaksi dari beberapa pelanggan yang memiliki korelasi yang sama, dapat dilihat pada tabel 4.

Tabel 4. Contoh Transaksi

No Nota	List Barang
S17082014-001	A, B, D, E
S07052014-013	D, E
S29092013-004	A, C, D
S03022011-041	A, B, D
S01052015-030	B

Dari contoh data transaksi yang memiliki kesamaan korelasi dicari large itemsetnya dengan minimum support 40%. Minimum support tersebut sama dengan minimum muncul pada 2 (dua) transaksi. Dari hasil pencarian tersebut didapatkan large itemset hingga $i=2$. Hasil dari pencarian large itemset tersebut dapat dilihat pada gambar 2.

<p>C1:</p> <table border="1" style="margin: auto;"> <tr><td>A</td><td>3</td></tr> <tr><td>B</td><td>3</td></tr> <tr><td>C</td><td>1</td></tr> <tr><td>D</td><td>4</td></tr> <tr><td>E</td><td>2</td></tr> </table>	A	3	B	3	C	1	D	4	E	2	<p>L1:</p> <table border="1" style="margin: auto;"> <tr><td>A</td><td>3</td></tr> <tr><td>B</td><td>3</td></tr> <tr><td>D</td><td>4</td></tr> <tr><td>E</td><td>2</td></tr> </table>	A	3	B	3	D	4	E	2		
A	3																				
B	3																				
C	1																				
D	4																				
E	2																				
A	3																				
B	3																				
D	4																				
E	2																				
<p>C2:</p> <table border="1" style="margin: auto;"> <tr><td>AB</td><td>2</td></tr> <tr><td>AD</td><td>3</td></tr> <tr><td>AE</td><td>1</td></tr> <tr><td>BD</td><td>2</td></tr> <tr><td>BE</td><td>0</td></tr> <tr><td>DE</td><td>2</td></tr> </table>	AB	2	AD	3	AE	1	BD	2	BE	0	DE	2	<p>L2:</p> <table border="1" style="margin: auto;"> <tr><td>AB</td><td>2</td></tr> <tr><td>AD</td><td>3</td></tr> <tr><td>BD</td><td>2</td></tr> <tr><td>DE</td><td>2</td></tr> </table>	AB	2	AD	3	BD	2	DE	2
AB	2																				
AD	3																				
AE	1																				
BD	2																				
BE	0																				
DE	2																				
AB	2																				
AD	3																				
BD	2																				
DE	2																				
<p>C3:</p> <table border="1" style="margin: auto;"> <tr><td>ABD</td><td>2</td></tr> </table>	ABD	2	<p>L3:</p> <table border="1" style="margin: auto;"> <tr><td>ABD</td><td>2</td></tr> </table>	ABD	2																
ABD	2																				
ABD	2																				

Gambar 2. Hasil Large Itemset

Dari gambar 2 tersebut didapatkan hingga large itemset-3. Untuk pembentukan kandidat C4 tidak ada karena hasil dari large itemset-3 hanya 1, sehingga dari situ proses pencarian large itemset berhenti.

Langkah berikutnya adalah mendapatkan strong association rules. Large itemset yang dapat digunakan adalah large itemset mulai dari large itemset-2. Large itemset-1 tidak dapat digunakan untuk menghasilkan strong association rule karena tidak ada RHS-nya. Dari contoh large itemset yang dihasilkan pada gambar 2, didapatkan strong association rule yang memenuhi minimum confidence 75% pada gambar 3.

CSAR:	
A → B ; conf = 66,7%	
B → A; conf = 66,7%	
A → D; conf = 100 %	
D → A; conf = 75%	
B → D; conf = 66,7%	
D → B; conf = 50%	
D → E; conf = 50%	
E → D; conf = 100%	
A → BD; conf = 66,7%	
B → AD; conf = 66,7%	
D → AB; conf = 50%	
	SAR:
	A → D; conf = 100 %
	D → A; conf = 75%
	E → D; conf = 100%

Gambar 3. Hasil Strong Association Rule

Pada gambar 3 ditunjukkan proses pembentukan kandidat strong association rule hingga strong association rule-nya, dimana confidence dari masing-masing kandidat yang tidak memenuhi minimum confidence akan disingkirkan. Hanya kandidat yang memenuhi minimum confidence yang akan menjadi strong association rules.

Strong Association Rule yang dihasilkan tersebut akan ditawarkan sebagai rekomendasi paket, dimana harga yang ditawarkan apabila membeli secara bersamaan adalah total harga barang tersebut dikurangi 5%. Sebagai contoh barang A Rp 10.000,- sedangkan barang D Rp 15.000,-. Saat pelanggan akan membeli barang A, maka diberikan rekomendasi barang D dengan total yang harus dibayar adalah Rp 23.750,-.

UJI COBA

Dari data transaksi yang diuji cobakan sebanyak 500 transaksi, dimana digunakan 80:20 untuk data training dan data testing. Dengan minimum support 40% dan minimum confidence 75% didapatkan bahwa terdapat 21 rekomendasi yang tidak sesuai. Dari percobaan tersebut dapat disimpulkan bahwa akurasi dari percobaan ini adalah 79%. Adapun pada table 5 diujicobakan minimum support dan minimum confidence yang berbeda.

Tabel 5. Uji Coba Akurasi

No.	Min sup	Min Conf	Rekomendasi Error
1	60	80	9
2	50	50	37
3	40	90	7

Dari tabel 5 dapat pula disimpulkan bahwa untuk mendapatkan rekomendasi yang baik, minimum confidence harus diberikan angka yang lebih tinggi.

KESIMPULAN

Setelah melakukan percobaan pada data transaksi didapatkan beberapa kesimpulan dari penelitian ini, yaitu:

1. Collaborative filtering dapat membantu menentukan profil pengguna dengan melihat korelasinya.
2. Link Analysis dapat membantu menentukan rekomendasi yang tepat sesuai dengan profil pengguna.
3. Semakin besar nilai minimum confidence, maka hasil strong association rule akan semakin baik.

DAFTAR PUSTAKA

- [1] Ampaziz, N. 2008. Collaborative Filtering via Concept Decomposition on the Netflix Dataset. ECAI Workshop on Recommender System.
- [2] Desrosiers, Christian & George Karypis, A comprehensive survey of neighborhood-based recommendation methods.
- [3] Jia Weihai, Association Analysis,
<http://www.cs.sfu.ca/CourseCentral/310/qyang/lectures/association.pdf>
- [4] Anita Wasilewska, Apriori Algorithm,
http://www3.cs.stonybrook.edu/~cse634/lecture_notes/07apriori.pdf
- [5] Rakesh Agrawal, Fast Algorithms for Mining Association Rules, <http://rakesh.agrawal-family.com/papers/vldb94apriori.pdf>

Halaman ini sengaja dikosongkan