

Market Basket Analysis Menggunakan Algoritme Apriori

Market Basket Analysis Using Apriori Algorithm

¹Faris Lailatul Ramdhan, ²Anneke Iswani Achmad, ³Aceng Komarudin Mutaqin

^{1,2,3}Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung,
Jl. Tamansari No.1 Bandung 40116

email: ¹farislailatulramdhan@gmail.com, ²annekeiswani11@gmail.com, ²aceng.k.mutaqin@gmail.com

Abstract. Market basket analysis is one method in data mining that aims to find products that are frequently purchased together and the association between products in the transaction data in a retail store. This analysis is constructed with 2-step that is looking for a set of products that are often purchased and found an association between products. One of the algorithms used to find the set of frequently purchased products is the a priori algorithms. While the association between products is sought based on two measures, i.e., support and confidence. Support between the two sets of products stating the percentage of the number of sales transactions are two sets of these products together in the overall transaction. While confidence between two sets of products stating the percentage of the number of sales transactions are the two sets of data together in a product throughout the sales transaction that contains one set of products. This paper discusses the market basket analysis that searches the set of products that often are bought using the a priori algorithms. The data used in this paper is the purchase transaction data for 1 year in one of the pastry shop named Extended Bakery in America. The data contains 20.000 transactions with 50 types of products. As a result, the Apriori algorithm obtained 50 frequent 1-itemsets, 26 frequent 2-itemsets, and 4 frequent 3-itemsets as well as products that have to be placed side by side is the Opera Cake, Tart Cherry and Apricot Danish.

Keywords: Market Basket Analysis, Data Mining, Apriori, Support, Confidence.

Abstrak. *Market basket analysis* merupakan salah satu metode dalam penambangan data (*data mining*) yang bertujuan untuk menemukan produk-produk yang sering dibeli bersamaan dan asosiasi antar produk dalam data transaksi di toko ritel. Analisis ini dibangun dengan 2 langkah yaitu mencari himpunan produk yang sering dibeli dan menemukan asosiasi antar produk. Salah satu algoritme yang digunakan untuk mencari himpunan produk yang sering dibeli adalah algoritme Apriori. Sedangkan asosiasi antar produk dicari berdasarkan dua ukuran, yaitu *support* dan *confidence*. *Support* antara dua himpunan produk menyatakan persentase banyaknya transaksi penjualan dua himpunan produk tersebut secara bersama-sama dalam keseluruhan transaksi. Sedangkan *confidence* antara dua himpunan produk menyatakan persentase banyaknya transaksi penjualan dua himpunan produk secara bersama-sama dalam data seluruh transaksi penjualan yang memuat salah satu himpunan produk tersebut. Makalah ini membahas *market basket analysis* yang pencarian himpunan produk yang sering dibelinya menggunakan algoritme apriori. Data yang digunakan dalam makalah ini adalah data transaksi pembelian selama 1 tahun di salah satu toko kue yang bernama *Extended Bakery* di Amerika. Data tersebut memuat 20.000 transaksi dengan 50 jenis produk. Hasilnya, dengan algoritme Apriori didapatkan 50 himpunan 1 produk, 26 himpunan 2 produk, dan 4 himpunan 3 produk serta produk yang harus diletakkan berdampingan adalah *Opera Cake*, *Cherry Tart* dan *Apricot Danish*.

Kata Kunci: Market Basket Analysis, Data Mining, Apriori, Support, Confidence.

A. Pendahuluan

Seiring dengan perkembangan teknologi, semakin berkembang pula kemampuan kita dalam mengumpulkan dan mengolah data. Penggunaan sistem komputerisasi dalam berbagai bidang baik itu dalam transaksi-transaksi bisnis, maupun untuk kalangan pemerintah dan sosial, telah menghasilkan data yang berukuran sangat besar. Data yang terkumpul ini merupakan suatu tambang emas yang dapat digunakan sebagai informasi dalam dunia bisnis.

Aplikasi basis data telah banyak diterapkan dalam berbagai bidang, antara lain bidang manajemen, manajemen data untuk industri, ilmu pengetahuan, administrasi pemerintah dan bidang-bidang lainnya. Akibatnya data yang dihasilkan oleh bidang-bidang tersebut sangatlah besar dan berkembang sangat cepat. Hal ini menyebabkan timbulnya kebutuhan terhadap teknik-teknik yang dapat melakukan pengolahan data sehingga dari data yang ada dapat diperoleh informasi penting yang dapat digunakan untuk perkembangan pada masing-masing bidang.

Perkembangan dunia bisnis ritel kembali diserbu konsumen yang ditandai dengan semakin banyaknya toko ritel yang didirikan. Hampir semua toko ritel menggunakan sistem komputerisasi dalam penyimpanan data penjualan dan pembelian, sehingga akan dihasilkan banyak data transaksi. Data transaksi tersebut dapat digunakan sebagai sumber informasi penting dalam mempertahankan keberadaan usaha tersebut (Lestari, 2009).

Salah satu upaya yang dilakukan untuk dapat mempertahankan bisnis ritel ini adalah dengan meningkatkan dan mempertahankan tingkat pertumbuhan penjualannya. Pendekatan yang dapat dilakukan dalam rangka meningkatkan volume penjualan tersebut adalah dengan memahami dan mengetahui pola perilaku belanja konsumen. Pengetahuan tentang pola perilaku belanja tersebut kemudian dapat dimanfaatkan untuk menyusun tata letak rak toko sedemikian rupa sehingga memberikan kemudahan bagi konsumen dalam melakukan pembelian, dan dapat meningkatkan kemungkinan munculnya keinginan membeli.

Pada sebuah toko ritel dapat digali berbagai macam informasi dari data transaksi pembelian produk-produk. Asosiasi antara jenis-jenis produk yang dibeli konsumen misalnya kapan seseorang membeli beberapa produk secara bersamaan (misalnya roti, susu dan mentega), dan kapan produk tertentu banyak dibeli, merupakan informasi yang dapat digunakan pihak toko ritel dalam mengatur peletakan produk di etalase dan rak produk, sehingga dapat membantu pelayan toko ritel dan konsumen dalam mencari dan mengingat kembali bila lupa. Pihak toko ritel dapat menentukan kapan perlu mendatangkan lebih banyak jenis produk tertentu pada waktu-waktu tertentu. Keputusan-keputusan ini pada akhirnya dapat membantu toko ritel untuk lebih mengoptimalkan kinerja.

Salah satu metode yang dapat digunakan untuk menganalisis pola perilaku belanja konsumen di toko ritel adalah *market basket analysis* (Annie dan Kumar, 2012). Analisis ini merupakan salah satu metode dalam penambangan data (*data mining*) yang bertujuan untuk menemukan produk-produk yang sering dibeli bersamaan dan asosiasi antar produk dalam data transaksi. Tujuan analisis ini pada transaksi penjualan adalah untuk merancang strategi penjualan atau pemasaran yang efektif dengan memanfaatkan data transaksi penjualan yang telah tersedia di toko ritel. Salah satu algoritme yang dapat digunakan dalam *market basket analysis* untuk menemukan produk-produk yang sering dibeli bersamaan dari data transaksi penjualan adalah algoritme Apriori (Han dkk., 2012). Algoritme Apriori adalah suatu algoritme yang paling terkenal untuk menemukan pola-pola penjualan produk di dalam suatu

database transaksi penjualan di toko ritel yang memiliki frekuensi penjualan di atas ambang batas yang ditentukan. Setelah ditemukan produk-produk yang sering dibeli bersamaan melalui algoritme Apriori, langkah selanjutnya dalam *market basket analysis* adalah menemukan asosiasi antar produk dalam data transaksi. Ada dua ukuran asosiasi yaitu *support* dan *confidence*. *Support* antara himpunan produk *A* dan himpunan produk *B* menyatakan persentase banyaknya transaksi penjualan himpunan produk *A* yang dibeli oleh konsumen bersama-sama dengan himpunan produk *B* dalam seluruh transaksi penjualan yang ada. *Confidence* antara himpunan produk *A* dan himpunan produk *B* menyatakan persentase banyaknya transaksi penjualan himpunan produk *A* yang dibeli oleh konsumen bersama-sama dengan himpunan produk *B* dalam seluruh transaksi penjualan yang ada himpunan produk *A*. Data yang akan digunakan dalam analisis ini adalah data transaksi penjualan toko *Extended Bakery* yang berlokasi di California, Amerika Serikat.

Berdasarkan uraian di atas, tujuan dari makalah ini adalah:

1. Untuk menemukan produk-produk yang sering dibeli bersamaan (*frequent itemsets*) dari data transaksi penjualan di toko *Extended Bakery* yang berlokasi di California, Amerika Serikat menggunakan algoritme Apriori.
2. Untuk mengetahui besar persentase banyaknya transaksi pembelian dua himpunan produk secara bersama-sama oleh konsumen dalam data transaksi penjualan (*support*) di toko *Extended Bakery* yang berlokasi di California, Amerika Serikat.

Untuk mengetahui besar persentase banyaknya transaksi penjualan dua himpunan produk secara bersama-sama oleh konsumen dalam data seluruh transaksi penjualan yang memuat salah satu himpunan produk tersebut (*confidence*) di toko *Extended Bakery* yang berlokasi di California, Amerika Serikat.

B. Landasan Teori

Market basket analysis adalah analisis dari kebiasaan membeli konsumen dengan mencari asosiasi dari produk-produk berbeda yang diletakkan konsumen dalam keranjang belanjanya. Tujuan dari *market basket analysis* adalah untuk menentukan produk-produk (jasa) apa saja yang paling sering dibeli atau digunakan sekaligus oleh para konsumen. Istilah analisis ini sendiri datang dari kejadian yang sudah sangat umum terjadi di dalam supermarket, yaitu ketika para konsumen memasukan semua produk yang mereka beli ke dalam keranjang yang umumnya telah disediakan oleh pihak toko ritel itu sendiri.

Menurut David Olson dan Yong Shi (2008), *market basket analysis* mengacu pada berbagai metodologi yang mempelajari komposisi keranjang belanja yang terdiri atas produk-produk yang dibeli pada satu kejadian belanja. Teknik ini telah dikembangkan dalam berbagai operasi pasar swalayan. Data keranjang belanja dalam bentuknya yang paling mentah adalah daftar transaksi pembelian oleh konsumen, yang mengindikasikan hanya produk-produk yang dibeli bersamaan. Data ini menantang karena beberapa hal:

1. Ukuran datanya besar, biasanya jutaan transaksi pertahun.
2. Setiap keranjang belanja hanya mengandung sebagian kecil dari produk yang dijual.
3. Orang-orang dengan selera yang berbeda cenderung sekelompok produk yang spesifik).

Albion Research (2007), mengemukakan informasi mengenai produk-produk yang biasanya dibeli secara bersama-sama oleh para konsumen dapat memberikan wawasan tersendiri bagi para pengelola toko atau swalayan untuk menaikkan laba bisnisnya. Misalnya, seorang manajer toko bisa saja memanfaatkan informasi seperti ini untuk menempatkan produk-produk yang umumnya dibeli bersama-sama ke dalam sebuah area yang berdekatan jaraknya. Hal ini bertujuan untuk memudahkan mayoritas pembeli agar semakin senang membeli beberapa produk berlainan sekaligus. Informasi ini juga dapat digunakan untuk menyampaikan ide mendasar dari *market basket analysis*, dimulai dengan melihat gambar keranjang belanja yang berisi bermacam-macam produk yang dibeli oleh seseorang di sebuah supermarket.

Keranjang ini berisi bermacam-macam produk seperti *orange juice*, pisang, *soft drink*, pembersih jendela, dan detergent. Sebuah keranjang memberitahukan kepada kita tentang apa saja yang dibeli oleh seorang konsumen dalam satu waktu. Sebuah daftar belanjaan yang lengkap yang diperoleh dari semua konsumen memberikan kita informasi yang sangat banyak, dan ini dapat menjelaskan produk apa saja yang paling penting dari bisnis penjualan yaitu jenis produk dan kapan konsumen membeli. Setiap konsumen membeli seperangkat produk yang berbeda, dalam jumlah yang berbeda dan dalam waktu yang berbeda. *Market basket analysis* menggunakan informasi apa yang dibeli oleh konsumen untuk menyediakan tanda atau informasi yaitu siapa mereka dan mengapa mereka melakukan pembelian tersebut.

Market basket analysis juga menyediakan pengertian tentang produk dengan memberitahukan kepada kita produk-produk mana yang memungkinkan untuk dibeli secara bersamaan dan produk mana yang lebih disetujui untuk dipromosikan. Informasi ini dapat digunakan dalam menentukan tata letak rak toko, menetapkan produk-produk mana yang akan diletakan dalam tempat yang khusus, untuk mengidentifikasi ketika persoalannya berpasangan, dan lain sebagainya. Teknik yang umum digunakan adalah *assosiation rules*. *Market basket analysis* mempunyai dua tahapan yaitu, menemukan himpunan produk yang sering dibeli (*frequent itemset*) dan membentuk *association rules*. Salah satu cara menemukan *frequent itemset* adalah dengan algoritme Apriori.

Algoritme Apriori adalah suatu algoritme yang diusulkan oleh Agrawal dan Srikant di tahun 1994 untuk menambang himpunan-himpunan produk yang sering dibeli (*frequent itemsets*). Istilah Apriori didasarkan pada fakta bahwa algoritme-nya menggunakan pengetahuan sebelumnya (*prior knowledge*) dari *frequent itemsets*. Algoritme Apriori merupakan suatu pendekatan iteratif, dimana himpunan-himpunan k produk (k -*itemsets*) digunakan untuk mencari himpunan-himpunan $k + 1$ produk ($(k + 1)$ -*itemsets*). Pertama, himpunan-himpunan 1 produk yang sering dibeli (*frequent 1-itemsets*) ditemukan dengan jalan mencari pada data transaksi, kemudian menghitung jumlah setiap produk dan mengumpulkan produk-produk itu yang memenuhi jumlah minimum himpunan produk (*minimum support count*). Hasilnya dinotasikan oleh L_1 . Selanjutnya L_1 digunakan untuk mencari L_2 yaitu himpunan-himpunan 2 produk yang sering dibeli (*frequent 2-itemsets*), yang digunakan untuk mencari L_3 , dan seterusnya sampai L_k .

Adapun Prosedur dari algoritme Apriori adalah sebagai berikut:

Input:

- D , data transaksi.
- Min_sup , batas jumlah minimum himpunan produk.

Output: L , *frequent itemsets* dalam data transaksi D .

Method:

```
(1)  $L_1$  = mencari frequent 1-itemsets( $D$ )
(2) for ( $k = 2; L_{k-1} \neq \emptyset; k++$ ) {
(3)    $C_k$  = apriori_gen( $L_{k-1}$ );
(4)   for each transaksi  $t \in D$  // mengecek transaksi yang ada daalm D
(5)      $C_t$  = subset( $C_k, t$ ); // mendapatkan himpunan produk yang jadi kandidat
(6)     for each kandidat  $c \in C_t$ 
(7)        $c.count++$ ;
(8)   }
(9)    $L_k = \{c \in C_t | c.count \geq min\_sup\}$ 
(10) }
(11) return  $L = \cup_k L_k$ ;
```

Prosedur apriori_gen (L_{k-1} : *frequent* ($k-1$) - *itemsets*)

```
(1) for each produk  $l_1 \in L_{k-1}$ 
(2)   for each produk  $l_2 \in L_{k-1}$ 
(3)     if ( $l_1[1] = l_2[1]) \wedge (l_1[2] = l_2[2]) \wedge \dots \wedge (l_1[k-2] = l_2[k-2]) \wedge (l_1[k-1] = l_2[k-1])$ 
(4)        $c = l_1 \bowtie l_2$ ; // tahap penggabungan: membangkitkan kandidat
(5)       if has_infrequent_subset ( $c, L_{k-1}$ ) then
(6)         delete  $c$ ; // tahap pemangkasan: menghilangkan kandidat
(7)       else add  $c$  to  $C_k$ ;
(8)     }
(9)   return  $C_k$ 
```

Prosedur has_infrequent_subset (c : *candidate* k - *itemset*);

```
 $L_{k-1}$ : frequent ( $k-1$ ) - itemsets; // menggunakan pengetahuan sebelumnya
(1) for each ( $k-1$ ) himpunan bagian  $s$  dari  $c$ 
(2)   if  $s \notin L_{k-1}$  then
(3)     return TRUE;
return FALSE.
```

Setelah menemukan seluruh *frequent itemset* yang mungkin, tahap selanjutnya adalah membentuk *association rules*.

Dalam *association rules* diperlukan ukuran yang ditentukan oleh pengguna untuk mengatur batasan sejauh mana dan sebanyak apa hasil output yang diinginkan. Ukuran tersebut adalah *support* dan *confidence*.

Support adalah persentase banyaknya transaksi yang memuat himpunan produk yang dibeli secara bersamaan dari keseluruhan transaksi. *Support* antara himpunan produk A dan himpunan produk B menyatakan persentase banyaknya transaksi penjualan himpunan produk A yang dibeli oleh konsumen bersama-sama dengan himpunan produk B dalam seluruh transaksi penjualan yang ada. Berikut ini adalah rumus untuk menghitung *support* himpunan produk A dan B .

$$Support(\{A, B\}) = P(\{A, B\})$$

$$= \frac{\text{banyaknya transaksi yang memuat } A \text{ dan } B}{\text{banyaknya transaksi}} \quad (2.1)$$

Misalnya $support(\{susu, roti\}) = 20\%$, ini memberikan informasi bahwa ada 20% transaksi pembelian susu dan roti secara bersamaan dari keseluruhan transaksi.

Confidence antara dua himpunan produk menyatakan persentase banyaknya transaksi penjualan kedua himpunan produk tersebut yang dibeli bersamaan oleh konsumen dari seluruh transaksi penjualan yang memuat salah satu dari dua himpunan produk di atas.

$Confidence(A \Rightarrow B)$ menyatakan persentase banyaknya transaksi penjualan himpunan produk A yang dibeli bersamaan oleh konsumen dengan himpunan produk B dari seluruh transaksi penjualan yang memuat himpunan produk A . Nilai $confidence$ menunjukkan ukuran kuatnya hubungan antar produk. Berikut ini adalah rumus untuk menghitung $confidence$ himpunan produk A dan B .

$$\begin{aligned} Confidence(A \Rightarrow B) &= P(B|A) \\ &= \frac{P(A \text{ dan } B)}{P(A)} \\ &= \frac{\text{banyaknya transaksi yang memuat } A \text{ dan } B}{\text{banyaknya transaksi yang memuat } A} \end{aligned} \quad (2.2)$$

Misalnya $confidence(susu \Rightarrow roti) = 50\%$, ini memberikan informasi bahwa ada 50% transaksi pembelian susu dan roti secara bersamaan dari keseluruhan transaksi yang memuat pembelian susu. Artinya ada 50% transaksi yang sebelumnya membeli susu kemudian membeli roti.

C. Hasil Penelitian dan Pembahasan

Berikut adalah hasil dari pembahasan dari penerapan algoritme Apriori dan *association rules* untuk data transaksi pembelian produk oleh konsumen di toko *Extended Bakery* yang berlokasi di *California*, Amerika Serikat. Lama pengamatan data ini adalah 1 tahun. Data tersebut memuat 20.000 transaksi dan 50 jenis produk. Data transaksi pembelian produk dibuat kedalam bentuk matriks terlebih dahulu. Kemudian algoritme Apriori dilakukan dengan bantuan Matlab 2008b.

Dalam bagian ini, banyak para ahli menggunakan *minimum support count* 3% dari total transaksi yang ada (3% dari 20.000 adalah 600 transaksi).

Tabel 1. Himpunan 1 Produk yang Sering Dibeli (L_1)

No Himpunan Produk	Himpunan Produk	Banyaknya Transaksi	No Himpunan Produk	Himpunan Produk	Banyaknya Transaksi
1	{1}	1.673	26	{26}	890
2	{2}	1.720	27	{27}	848
3	{3}	1.506	28	{28}	1.729
4	{4}	1.673	29	{29}	1.973
5	{5}	1.840	30	{30}	1.390
6	{6}	1.693	31	{31}	841
7	{7}	852	32	{32}	1.420
8	{8}	2.197	33	{33}	1.634
9	{9}	854	34	{34}	1.634
10	{10}	1.690	35	{35}	892
11	{11}	811	36	{36}	1.854
12	{12}	1.483	37	{37}	1.351
13	{13}	1.385	38	{38}	1.463
14	{14}	855	39	{39}	885
15	{15}	1.686	40	{40}	823
16	{16}	1.534	41	{41}	1.331
17	{17}	1.678	42	{42}	1.369
18	{18}	1.527	43	{43}	1.848
19	{19}	1.825	44	{44}	1.243
20	{20}	1.519	45	{45}	1.443
21	{21}	831	46	{46}	2.072
22	{22}	866	47	{47}	1.645
23	{23}	1.837	48	{48}	1.535
24	{24}	1.389	49	{49}	1.306
25	{25}	1.365	50	{50}	1.422

Sumber: Hasil Pengolahan Matlab R2008b

Berdasarkan Tabel 1, dapat diketahui bahwa dari 50 kandidat himpunan 1 produk (C_1) yang ada semuanya memenuhi *minimum support count* 600 transaksi, disimpan di himpunan 1 produk yang sering dibeli (L_1). Informasi himpunan 1 produk yang ada dalam Tabel 1, digunakan untuk membangkitkan kandidat himpunan 2 produk (C_2). Didapatkan 1.225 himpunan 2 produk.

Tabel 2. Himpunan 2 Produk yang Sering Dibeli (L_2)

No Himpunan Produk	Himpunan Produk	Banyaknya Transaksi	No Himpunan Produk	Himpunan Produk	Banyaknya Transaksi
1	{1,3}	717	14	{12,46}	617
2	{1,47}	881	15	{15,45}	714
3	{2,20}	740	16	{16,50}	603
4	{3,47}	714	17	{17,33}	837
5	{4,19}	873	18	{17,46}	714
6	{4,36}	867	19	{18,30}	611
7	{5,10}	891	20	{18,48}	735
8	{6,23}	867	21	{19,36}	1.051
9	{8,12}	745	22	{28,29}	971
10	{8,16}	735	23	{30,48}	619
11	{8,38}	725	24	{33,46}	702
12	{8,46}	634	25	{34,43}	878
13	{12,38}	719	26	{38,46}	617

Sumber: Hasil Pengolahan Matlab R2008b

Berdasarkan Tabel 2, dapat diketahui bahwa dari 1.225 kandidat himpunan 2 produk (C_2), ada 26 himpunan produk yang memenuhi *minimum support count* 600, disimpan di himpunan 2 produk yang sering dibeli (L_2). Informasi himpunan 2 produk yang ada dalam Tabel 2, digunakan untuk membangkitkan kandidat himpunan 3 produk (C_3). Didapatkan 11 himpunan 3 produk.

Tabel 3. Himpunan 3 Produk yang Sering Dibeli (L_3)

No Himpunan Produk	Himpunan Produk	Banyaknya Transaksi
1	{1,3,47}	678
2	{4,19,36}	820
3	{8,12,38}	683
4	{17,33,46}	652

Sumber: Hasil Pengolahan dengan Matlab R2008b

Berdasarkan Tabel 3, dapat diketahui bahwa dari 11 kandidat himpunan 3 produk (C_3), ada 4 himpunan produk yang memenuhi *minimum support count* 600, disimpan di himpunan 3 produk yang sering dibeli (L_3). Berdasarkan Tabel 3. Terlihat bahwa tidak ada minimal dua himpunan produk yang produk 1 dan 2-nya sama. Dengan demikian tidak ada kandidat himpunan 4 produk yang dapat dibuat (C_4), oleh karena itu algoritme Apriori dihentikan. Setelah ditemukan semua himpunan produk yang mungkin, kemudian cari nilai *support* dan *confidence*.

Tabel 4. Nilai *Support* Himpunan 2 Produk

Himpunan Produk	Banyaknya Transaksi	<i>Support</i> (%)	Himpunan Produk	Banyaknya Transaksi	<i>Support</i> (%)
{19,36}	1051	5,26	{8,38}	725	3,63
{28,29}	971	4,86	{12,38}	719	3,60
{5,10}	891	4,46	{1,3}	717	3,59
{1,47}	881	4,41	{3,47}	714	3,57
{34,43}	878	4,39	{15,45}	714	3,57
{4,19}	873	4,37	{17,46}	714	3,57
{4,36}	867	4,34	{33,46}	702	3,51
{6,23}	867	4,34	{18,48}	634	3,17
{17,33}	837	4,19	{8,46}	619	3,10
{8,12}	745	3,73	{30,48}	617	3,09
{2,20}	740	3,70	{38,46}	617	3,09
{8,16}	735	3,68	{18,30}	611	3,06
{18,48}	735	3,68	{16,50}	603	3,02

Sumber: Hasil Pengolahan Matlab R2008

Berdasarkan Tabel 4. Menunjukkan bahwa dari total 20.000 transaksi, ada sekitar 5,26% transaksi yang memuat produk makanan *Cherry Tart* dan *Apricot Danish* yang dibeli bersamaan oleh konsumen. Begitu seterusnya untuk himpunan 2 produk yang ada pada tabel 4.

Tabel 5. Nilai *Support* Himpunan 3 Produk

Himpunan Produk	Banyaknya Transaksi	<i>Support</i> (dalam %)
{4,19,36}	820	4,10
{8,12,38}	683	3,42
{1,3,47}	678	3,39
{17,33,46}	652	3,26

Sumber: Hasil Pengolahan Matlab R2008b

Berdasarkan Tabel 5. Menunjukkan bahwa dari total 20.000 transaksi, ada sekitar 4,10% transaksi yang memuat produk makanan *Opera Cake*, *Cherry Tart* dan *Apricot Danish* yang dibeli bersamaan oleh konsumen. Begitu seterusnya untuk himpunan 3 produk yang ada pada tabel 5. Selanjutnya dicari nilai *confidence*-nya untuk setiap himpunan produknya.

Tabel 6. Nilai *Confidence* untuk Himpunan 2 Produk

Himpunan Produk	Nilai <i>Confidence</i> ($A \Rightarrow B$) (dalam %)	Nilai <i>Confidence</i> ($B \Rightarrow A$) (dalam %)
{1,3}	42,86	47,61
{1,47}	52,66	53,56
{2,20}	43,02	48,72
{3,47}	47,41	43,40
{4,19}	52,18	47,84
{4,36}	51,82	39,46
{5,10}	48,42	52,72
{6,23}	51,21	47,20
{8,12}	33,91	50,24
{8,16}	33,45	49,56
{8,38}	33,00	49,56
{8,46}	28,86	30,60
{12,38}	48,48	49,15
{12,46}	44,55	29,78
{15,45}	42,35	49,48
{16,50}	39,31	39,49
{17,33}	49,88	51,22
{17,46}	46,76	34,46
{18,30}	40,01	43,96
{18,48}	40,27	47,88
{19,36}	57,59	57,21
{28,29}	56,16	49,21
{30,48}	44,53	40,33
{33,46}	42,96	33,88
{34,43}	53,73	47,51
{38,46}	42,17	29,78

Sumber: Hasil Pengolahan Matlab R2008b

Berdasarkan Tabel 6. terlihat bahwa nilai *confidence* terbesar untuk himpunan 2 produk adalah nilai *Confidence* ($19 \Rightarrow 36$), yaitu ada sebesar 57,59% transaksi yang sebelumnya membeli produk *Cherry Tart* kemudian membeli produk *Apricot Danish*.

Tabel 7. *Confidence* Himpunan 3 Produk

Himpunan Produk {A, B, C}	Confidence (A \Rightarrow {B, C}) (dalam %)	Confidence (B \Rightarrow {A, C}) (dalam %)	Confidence (C \Rightarrow {A, B}) (dalam %)	Confidence ({B, C} \Rightarrow A) (dalam %)	Confidence ({A, C} \Rightarrow B) (dalam %)	Confidence ({A, B} \Rightarrow C) (dalam %)
{1,3,47}	40,53	45,02	41,22	76,96	94,56	94,96
{4,19,36}	49,01	44,93	44,23	94,58	93,93	78,02
{8,12,38}	31,09	46,06	46,68	94,21	91,68	94,99
{17,33,4}	38,86	39,90	31,47	91,32	77,90	92,88

Sumber: Hasil Pengolahan Matlab R2008b

Berdasarkan Tabel 7. terlihat bahwa nilai *confidence* terbesar untuk himpunan 3 produk adalah nilai *Confidence* ($\{8, 12\} \Rightarrow 38$), yaitu ada sebesar 94,99% transaksi yang sebelumnya membeli produk *Coffee Éclair* dan *Apple Pie* kemudian membeli produk *Almond Twist*.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Aplikasi algoritme Apriori dengan *minimum support count* 3% dari keseluruhan transaksi didapatkan 50 himpunan 1 produk yang sering dibeli, 26 himpunan 2 produk yang sering dibeli dan 4 himpunan 3 produk yang sering dibeli. Algoritme dihentikan di himpunan 3 produk karena tidak ada minimal dua himpunan produk yang produk 1 dan 2-nya sama.
2. Untuk himpunan 2 produk, produk makanan *Cherry Tart* dan *Apricot Danish* harus diletakkan berdampingan di etalase penjualan produk, karena paling sering dibeli bersamaan oleh konsumen, dengan persentase jumlah transaksi sebesar 5,26%. Untuk himpunan 3 produk, produk makanan *Opera Cake*, *Cherry Tart* dan *Apricot Danish* harus diletakkan berdampingan di etalase penjualan produk, karena paling sering dibeli bersamaan oleh konsumen, dengan persentase jumlah transaksi sebesar 4,10%.
3. Untuk himpunan 2 produk, ada 57,59% transaksi yang sebelumnya membeli produk makanan *Cherry Tart* kemudian membeli produk makanan *Apricot Danish*. Untuk himpunan 3 produk, ada 94,99% transaksi yang sebelumnya membeli produk makanan *Coffee Éclair* dan *Apple Pie* kemudian membeli produk makanan *Almond Twist*.

E. Saran

Saran yang dapat dikemukakan dalam makalah ini adalah:

1. Disarankan kepada pemilik toko *Extended Bakery* untuk menata ulang penempatan produk di etalase sesuai dengan hasil analisis ini yang berkaitan dengan produk-produk yang sering dibeli bersamaan.
2. Disarankan kepada peneliti lain untuk meneliti *Market Basket Analysis* dengan algoritme yang berbeda seperti algoritme *Eclat*, dan algoritme *FP-Growth*.

Daftar Pustaka

- Agrawal, R., dan Srikant, R. (1994). Fast algorithms for mining association rules. *Proceedings of the 20th VLDB conference*, 487–499.
- Albion Research Ltd. (2007). Market Basket Analysis. http://www.albionresearch.com/data_mining/market_basket.php. Diunduh tanggal 17 Febuari 2016.
- Annie, M. C., L. C., dan Kumar, D., A. (2012). Market Basket Analysis for a Supermarket based on Frequent Itemset Mining . *International Journal of Computer Science Issues*, 257-264.
- Dekhlyar, A. (2009). Developing an Undergraduate Course in Knowledge Discovery from Data, poster, 12th CSURegional Symposium on University Teaching, Cal Poly, San Luis Obispo. Diunduh tanggal 27 Juni 2016
- Giudici, P. (2009): *Applied data mining for business and industry*, Wiley, London.
- Goharian, N., dan Grossman, D. (2003). Introduction to Data Mining. <http://www.ir.iit.edu/~nazli/cs422/CS422-Slides/DMintroduction.pdf>. Diunduh tanggal 20 Februari 2016.
- Han, J., Kamber, M., dan Pei, J. (2012). *Data Mining: Concept and Techniques, Third Edition*. Waltham: Morgan Kaufmann Publishers.
- Hornick, L. A. (2006). *SThe Evolution of Product Placement: Consumer Awareness*

and. Ethical Cosiderations, West Virginia University.

Lestari, T. (2009). Analisis Keranjang Belanja Pada Data Transaksi Penjualan. Skripsi pada Fakultas Ekonomi IPB Bogor: Diterbitkan.

Olson, D., dan Shi, Y. (2008). *Pengantar Ilmu Penggalian Data Bisnis*. Salemba Empat, Jakarta.