

Market Basket Analysis Menggunakan Algoritma Eclat (Kasus Data Transaksi Penjualan Groceries)

Market Basket Analysis Using Eclat Algorithms (Case of Groceries Sales Transaction Data)

¹Taufik Qurrahman, ²Anneke Iswani Achmad

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

email: ¹qurrahman27@gmail.com, ²annekeiswani11@gmail.com

Abstract. Data mining is the mining or discovery of new information by looking for certain patterns or rules from a very large amount of data, its use has also been widely applied in various fields. One of the data mining techniques is Market Basket Analysis which is used to find associations between different sets of products that customers put in a basket. This thesis will discuss Market Basketball Analysis using the Equivalence Class Transformation (Eclat) algorithm on the sales transaction data of groceries to find associations among products that are often purchased simultaneously measured through a support value that shows how much opportunity the product purchase transactions contain itemsets is. purchased simultaneously, confidence value is a measure that shows the relationship between 2 items in a conditional manner (for example, counting likely how often item B is purchased by the customer if the customer buys an item A). Of 9,835 product sales transactions consisting of 169 types of products. The result, with Eclat algorithm obtained 59 frequent 1-Itemsets, 61 frequent 2-Itemsets and 2 Frequent 3-Itemsets, where the largest support value for 2-Itemsets was 0.0748 in other vegetables and whole milk products, which of the product confidence values first to be offered to consumers is that other vegetables after that are whole milk products. Whereas for the biggest support value for 3-Itemsets is 0.0232 on root vegetables, other vegetables and whole milk products, where from the confidence value of the product that must be offered first to consumers are other vegetables and whole milk, then the root vegetables.

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

Abstrak. *Data mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar, pemanfaatannya pun sudah banyak diterapkan dalam berbagai bidang. Salah satu teknik data pada *data mining* yaitu *Market Basket Analysis* yang digunakan untuk menemukan asosiasi diantara himpunan produk yang berbeda yang diletakkan pelanggan dalam keranjang. Skripsi ini akan membahas *Market Basket Analysis* menggunakan algoritma *Equivalence Class Transformation (Eclat)* pada data transaksi penjualan *groceries* untuk menemukan asosiasi diantara produk-produk yang sering dibeli secara bersamaan diukur melalui nilai *support* yaitu ukuran yang menunjukkan seberapa besar peluang banyaknya transaksi pembelian produk yang memuat *itemsets* yang dibeli secara bersamaan, nilai *confidence* yaitu suatu ukuran yang menunjukkan hubungan antara 2 *item* secara *conditional* (misal, menghitung kemungkinan seberapa sering *item B* dibeli oleh pelanggan jika pelanggan tersebut membeli sebuah *item A*). Dari 9.835 transaksi penjualan produk yang terdiri dari 169 jenis produk. Hasilnya, dengan algoritma *Eclat* didapatkan 59 *frequent 1-Itemsets*, 61 *frequent 2-Itemsets* dan 2 *Frequent 3-Itemsets*, dimana nilai *support* terbesar untuk *2-Itemsets* adalah 0,0748 pada produk *other vegetables* dan *whole milk*, dimana dari nilai *confidence* produk yang harus disodorkan terlebih dahulu kepada konsumen adalah *other vegetables* setelah itu barulah produk *whole milk*. Sedangkan untuk nilai *support* terbesar untuk *3-Itemsets* adalah 0,0232 pada produk *root vegetables*, *other vegetables* dan *whole milk*, dimana dari nilai *confidence* produk yang harus disodorkan terlebih dahulu kepada konsumen adalah *other vegetables* dan *whole milk*, setelah itu barulah produk *root vegetables*.

Kata Kunci: Market Basket Analysis, Data Mining, Eclat, Support, Confidence

A. Pendahuluan

Dalam dunia bisnis, khususnya industri swalayan, persaingan antar swalayan memasarkan tiap produk tidak bisa dilepaskan dari pemanfaatan teknologi informasi. Salah satu sumber informasi yang bisa digunakan untuk membantu kegiatan penjualan adalah sistem database. Dari database tersebut kita bisa menemukan suatu strategi yang tepat untuk mempertahankan bisnis ritel ini dengan berbagai persaingan yaitu dengan memanfaatkan informasi yang sebenarnya diinginkan oleh konsumen

untuk memberikan kemudahan untuk memilih barang belanjaan yang diinginkan. Sebagai contoh dalam peletakan barang-barang belanjaan yang tersusun di dalam rak sebaiknya disesuaikan dengan pola belanja konsumen untuk memudahkan konsumen mencari barang-barang yang diinginkan.

Pada bisnis ritel yang melibatkan salah satunya industri swalayan ini dapat digali berbagai macam informasi dari data transaksi penjualan produk produk guna membantu mengambil keputusan, sehingga hal ini mendorong munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola yang penting dan menarik dari data dalam jumlah besar, yang disebut dengan data mining. Penggunaan teknik data mining diharapkan dapat memberikan pengetahuan-pengetahuan yang sebelumnya tersembunyi di dalam gudang data sehingga menjadi informasi yang berharga.

Dimana seperti kita ketahui, setiap hari data transaksi penjualan di industri swalayan semakin bertambah banyak sehingga data tersebut menumpuk, namun data ini seringkali diperlakukan hanya sebagai rekaman tanpa pengolahan lebih lanjut sehingga tidak mempunyai nilai guna lebih

untuk bisa dimanfaatkan dengan baik. Analisis dari tiap koleksi data tersebut akan menghasilkan pengetahuan atau informasi, misalnya berupa pola dan kaidah asosiasi yang terjadi pada data.

Metode yang sering digunakan untuk melakukan analisis pola perilaku belanja konsumen adalah analisis keranjang belanja atau Market Basket Analysis (MBA). Analisis ini merupakan salah satu metode dalam penambangan data (data mining) yang mempelajari tentang perilaku kebiasaan konsumen dalam membeli barang secara bersamaan dalam satu waktu. Teknik ini digunakan untuk merancang suatu strategi penjualan dan pemasaran barang melalui proses pencarian asosiasi atau hubungan antar item data. Pencarian asosiasi berawal dari pengolahan data transaksi pembelian barang dari setiap pembeli, kemudian dicari hubungan antar barang-barang yang dibeli. Pencarian informasi ini hampir sama dengan mencari peluang kemunculan barang yang dibeli sesuai dengan kebiasaan berbelanja masyarakat dan jumlah transaksi yang ada.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut:

Berapakah besar nilai *support* berdasarkan *frequent itemsets* dari data *Groceries* dengan Algoritma *Eclat*?

Berapakah besar nilai *confidence* berdasarkan *frequent itemsets* dari data *Groceries* dengan Algoritma *Eclat*?

B. Landasan Teori

Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi

dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. (Turban, dkk. 2005). Data mining juga disebut sebagai analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data.

Sehingga dari beberapa pengertian yang dikemukakan oleh para pakar ilmuwan diatas, dapat disimpulkan bahwa data mining adalah suatu teknik untuk mengolah informasi yang jumlahnya sangat besar sehingga dari informasi tersebut nantinya akan didapatkan suatu data yang bermanfaat dan bisa digunakan.

Market Basket Analysis

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 memasukkan semua produk yang mereka beli ke dalam keranjang yang umumnya telah disediakan oleh pihak supermarket itu sendiri.

Market basket analysis mempunyai dua tahapan yaitu, menemukan himpunan produk yang sering dipinjam (*frequent itemset*) dan membentuk *association rules*. Salah satu cara menemukan *frequent itemset* adalah dengan algoritma *Eclat* (*Equivalence Class Transformation*). Masalah utama di dalam menemukan *itemset*, misalnya yang terkandung di dalam transaksi adalah terdapat banyak

sekali kemungkinan *itemset* yang muncul, yang membuat pendekatan naif menjadi tidak layak karena waktu eksekusi yang lama (Goethals, 2003). Namun, ada pendekatan yang bisa menjadi alternatif dan lebih canggih dengan dua algoritma dikenal dengan nama *Apriori* dan *Eclat* yang paling populer. Keduanya mengandalkan pencarian *top-down* dalam kisi subset dari item.

Algoritma Eclat (Equivalence Class Transformation)

Algoritma *Eclat* melakukan pencarian *frequent itemset* dari kumpulan data (dataset). Yang mana *frequent itemset* merupakan himpunan data yang paling sering muncul. Algoritma *Eclat* diperkenalkan oleh Zaki, Parthasarathy, Ogihara, & Li (1997) yaitu algoritma yang melakukan kegiatan pengelompokan item yang sama berdasarkan kriteria tertentu ke dalam kelas (*Equivalence class*). Kelas yang sama tersebut didapatkan dari partisi suatu himpunan. *Equivalence class* pada algoritma *Eclat* dibangun dengan kelas berbasis *prefix*. Algoritma *Eclat* mempunyai proses lebih cepat, karena dataset akan dipresentasikan dalam *vertical* format dari dataset. Menurut Li, Liu, Liao, & Choudhary (2006) algoritma *Eclat* akan mengubah bentuk dataset *horizontal* menjadi *vertikal* transaksi id-list (tidlist) dari *itemset*. Tid-list (TID) dari *itemset* diperoleh dari urutan ID data transaksi yang memuat *itemset* tersebut. Sehingga hal itulah yang membuat algoritma *eclat* berbeda dengan algoritma lainnya.

Association Rules Mining

Association rules mining adalah metode *data mining* yang menggunakan hubungan suatu *item* dengan *item* lainnya yang terdapat di dalam *itemset* untuk memprediksi pola dari kumpulan data (Olson & Delen, 2008).

Dalam *association rules* diperoleh *frequent itemsets* dari penentuan *minimum support* yang biasanya 2% dari keseluruhan transaksi yang digunakan untuk menyeleksi data, dari data yang sudah diseleksi akan dihitung *support* dan *confidence*.

Support adalah ukuran yang menunjukkan seberapa besar peluang banyaknya transaksi yang memuat *itemsets* yang dibeli secara bersamaan dari keseluruhan transaksi. Ukuran ini akan menentukan apakah suatu *itemsets* dapat dicari nilai *confidence*. Berikut ini adalah rumus untuk menghitung *support* produk *X* dan produk *Y*.

$$\begin{aligned}
 \text{Support}(\{X, Y\}) &= P(\{X, Y\}) \\
 &= \frac{n(X) \cap n(Y)}{n(S)} \\
 &= \frac{\text{banyaknya transaksi yang memuat } X \text{ dan } Y}{\text{total banyaknya transaksi}}
 \end{aligned}
 \tag{2.1}$$

Keterangan :

$$\begin{aligned}
 n(X) \cap n(Y) &= \text{banyaknya transaksi yang memuat } X \text{ dan } Y \\
 n(S) &= \text{total banyaknya transaksi}
 \end{aligned}$$

Misalnya

support ({produk *x*, produk *y*}) = 0.4, ini memberikan informasi bahwa peluang pembelian produk *x* dan produk *y* akan dibeli secara bersamaan dari keseluruhan data transaksi yaitu 0.4.

Confidence adalah suatu ukuran yang menunjukkan hubungan antara 2 *item* secara *conditional* (misal, menghitung kemungkinan seberapa sering *item* *Y* dibeli oleh pelanggan jika pelanggan tersebut membeli sebuah *item* *X*). Kedua ukuran ini nantinya berguna dalam menentukan kekuatan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua parameter tersebut yang ditentukan oleh pengguna. Bila suatu pola memenuhi kedua nilai minimum parameter yang sudah ditentukan sebelumnya, maka pola tersebut dapat disebut sebagai

interesting rule atau *strong rule*. Berikut ini adalah rumus untuk menghitung *confidence* produk *X* dan produk *Y*.

$$\begin{aligned}
 \text{Confidence}(X \Rightarrow Y) &= P(Y|X) \\
 &= \frac{P(X \text{ dan } Y)}{P(X)} \\
 &= \frac{\text{banyaknya transaksi yang memuat } X \text{ dan } Y}{\text{banyaknya transaksi yang memuat } X}
 \end{aligned}
 \tag{2.2}$$

C. Hasil Penelitian dan Pembahasan

Bab ini berisikan tentang hasil dan pembahasan dari penerapan algoritma *Eclat* untuk data transaksi penjualan *groceries*. Data transaksi ini berukuran 9835 x 169 Kemudian algoritma *Eclat* dilakukan dengan bantuan *software* RStudio-1.1.447. Algoritma *Eclat* digunakan untuk menemukan produk-produk yang sering dibeli oleh konsumen dengan menghitung nilai *support*, dan nilai *confidence* berdasarkan *frequent itemsets*. Data tersebut memuat 9.835 transaksi dimana ada 169 jenis produk-produk yang dibeli oleh konsumen yang disajikan dalam tabel 1

Tabel 1 Data Transaksi Pembelian Produk di toko *Groceries*

Transaksi	Produk yang dibeli
1	citrus fruit, semi-finished bread, margarine, ready soups
2	tropical fruit, yogurt, coffee
3	whole milk
4	pip fruit, yogurt, cream cheese, meat spreads
5	other vegetables, whole milk, condensed milk, long life, bakery product
⋮	⋮
9835	Chicken, tropical fruit, other vegetables, vinegar, shopping bags

Sumber: Lantz (2015)

Dari data transaksi pembelian produk-produk oleh konsumen di toko *groceries* akan dicari 1 produk terlebih

Terlihat bahwa produk yang sering dibeli adalah *whole milk* dengan jumlah transaksi 2.513.

Sedangkan Produk yang jarang dibeli oleh konsumen adalah *baby food*

Tabel 2. Frequent 1-Itemsets

No	Itemsets	Banyaknya Transaksi	No	Itemsets	Banyaknya Transaksi
1	{whole milk}	2513	31	{chicken}	422
2	{other vegetables}	1903	32	{white bread}	414
3	{rolls/buns}	1809	33	{cream cheese }	390
4	{soda}	1715	34	{waffles}	378
5	{yogurt}	1372	35	{salty snack}	372
6	{bottled water}	1087	36	{long life bakery product}	368
7	{root vegetables}	1072	37	{dessert}	365
8	{tropical fruit}	1032	38	{sugar}	333
9	{shopping bags}	969	39	{UHT-milk}	329
10	{sausage}	924	40	{hamburger meat}	327
11	{pastry}	875	41	{berries}	327
12	{citrus fruit}	814	42	{hygiene articles}	324
13	{bottled beer}	792	43	{onions}	305
14	{newspapers}	785	44	{specialty chocolate}	299
15	{canned beer}	764	45	{candy}	294
16	{pip fruit}	744	46	{frozen meals}	279
17	{fruit/vegetable juice}	711	47	{misc. beverages}	279
18	{whipped/sour cream}	705	48	{oil}	276
19	{brown bread}	638	49	{butter milk}	275
20	{domestic eggs}	624	50	{specialty bar}	269
21	{frankfurter}	580	51	{ham}	256
22	{margarine}	576	52	{beverages}	256
23	{coffee}	571	53	{meat}	254
24	{pork}	567	54	{ice cream}	246
25	{butter}	545	55	{sliced cheese}	241
26	{curd}	524	56	{hard cheese}	241
27	{beef}	516	57	{cat food}	229
28	{napkins}	515	58	{grapes}	220
29	{chocolate}	488	59	{chewing gum}	207
30	{frozen vegetables}	473			

dahulu yang dibeli oleh konsumen dan akan dihitung banyaknya transaksi yang membeli 1 produk tersebut. Dengan bantuan perangkat lunak RStudio-1.1.447 diperoleh himpunan produk serta banyaknya transaksi yang membeli produk tersebut.

Berdasarkan pengolahan pertama dapat diketahui produk yang sering dibeli oleh konsumen juga produk yang jarang dibeli konsumen.

dan *sound storage medium* dengan jumlah transaksi 1..Langkah selanjutnya adalah membuat tabel *frequent 1-itemset* yang banyak transaksinya lebih dari atau sama dengan *minimum support*, 2% dari keseluruhan transaksi atau sebesar 197 transaksi.

Seperti yang dijelaskan pada Tabel 2. dimana hasilnya terdapat 59 jenis produk dari pengolahan kedua

yang ada dalam Tabel 2 karena memenuhi banyaknya transaksi lebih dari 2%.

Informasi yang didapat pada Tabel 2 akan digunakan untuk mencari *freequent 2-Itemsets*. Hasil *freequent 2-Itemsets* akan disajikan pada tabel 3 dengan bantuan perangkat lunak RStudio-1.1.447 sebagai berikut.

Berdasarkan Tabel 4.3 terlihat bahwa dari dalam Tabel 2 terdapat 61 jenis kombinasi yang memuat *Frequent 2-Itemsets* yang banyak transaksinya lebih dari atau sama dengan *minimum support*, 2% dari keseluruhan transaksi atau sebesar 197 transaksi. Dalam Tabel 3 terlihat bahwa ada 736 transaksi yang memuat 2 produk *other vegetables* dan *whole milk* yang dibeli secara bersama-sama, ada 557 transaksi yang memuat 2 produk *whole milk* dan *rolls/buns* yang dibeli secara bersama-sama, ada 551 transaksi yang memuat 2 produk *whole milk* dan *yogurt* yang dibeli secara bersama-sama, ada 481 transaksi yang memuat 2 produk *root vegetables* dan *whole milk* yang dibeli secara bersama-sama, dan seterusnya. Berdasarkan Tabel 3 juga dapat diketahui pasangan produk yang sering dibeli oleh konsumen dan jarang dibeli oleh konsumen. Terlihat bahwa pasangan produk yang sering dibeli oleh konsumen adalah *other vegetables* dan *whole milk* dengan jumlah transaksi 736. Pasangan produk yang jarang dibeli oleh konsumen adalah *other vegetables* dan *butter* dengan jumlah transaksi 197. Langkah selanjutnya adalah membuat tabel *Frequent 3-Itemsets* yang banyak transaksinya lebih dari atau sama dengan *minimum support*, 2% dari keseluruhan transaksi atau sebesar 197 transaksi. Dari pengolahan terakhir terlihat bahwa ada 2 kombinasi 3 produk yang banyak transaksinya lebih dari atau sama dengan 2% dari keseluruhan transaksi

atau sebesar 197 transaksi. *Minimum support* 2% dari keseluruhan transaksi atau sebesar 197 transaksi, *Frequent Itemsets* yang dihasilkan hanya bisa sampai *Frequent 3-Itemsets* tidak bisa lebih. Oleh karena itu algoritma *ecolat* dihentikan.

Hasil Perhitungan Support

Pada bagian ini akan diuraikan perhitungan nilai *support*, yaitu ukuran yang menunjukkan seberapa besar peluang banyaknya transaksi yang memuat *itemsets* yang dibeli secara bersamaan dari keseluruhan data transaksi pembelian produk-produk di toko *Groceries*. Nilai *support* dari *itemsets* yang akan digunakan pada tahapan ini adalah *minimum support* = 0.02 yang merupakan nilai yang umum digunakan. Tabel 5 berisikan nilai *support* untuk seluruh *freequent 2-itemsets*. Sebagai contoh untuk menghitung nilai *support* pada *freequent 2-itemsets* pembelian produk *other vegetables* dan *whole milk* yaitu **0.07483** Nilai ini menunjukkan bahwa dari total 9.835 transaksi pembelian produk, peluang munculnya produk

other vegetables dan whole milk dibeli secara bersamaan yaitu sebesar 0.07483. Dengan cara yang sama, dapat dihitung nilai *support* untuk setiap *freequent 2-itemsets* lainnya. Nilai *support* ini disajikan dalam Tabel 5 dalam bentuk yang sudah diurut dari nilai *support* terbesar sampai terkecil. nilai *support* yang ada menjadi informasi bagi pemilik toko *Groceries* dalam hal penempatan produk yang dijualnya. Sebagai contoh, produk *other vegetables* dan *whole milk* harus diletakkan berdampingan di etalase/rak-rak penjualan produk, karena paling sering dibeli secara bersamaan oleh konsumen, dengan peluang sebesar 0,07483. Produk *whole milk* dan *rolls/buns* harus diletakkan berdampingan di etalase/rak

penjualan produk, karena berada di urutan ke-2 yang sering dibeli bersamaan oleh konsumen, dengan peluang sebesar 0,05663. Produk *whole milk* dan *yogurt* harus diletakkan berdampingan di etalase/rak penjualan produk, karena berada di urutan ke-3 yang sering dibeli bersamaan oleh konsumen, dengan peluang sebesar 0,05602. Begitu seterusnya untuk *frequent 2-itemsets* yang ada pada Tabel 5. Selanjutnya akan dihitung nilai *support* untuk *frequent 3-itemsets* yang dibeli oleh konsumen secara bersama-sama. Sebagai contoh *frequent 3-itemsets* *root vegetables, other vegetables* dan *whole milk* nilai *support*-nya dihitung menggunakan persamaan (2.1). Dengan cara yang sama dapat dihitung nilai *support* untuk *frequent 3-itemsets* lainnya yang disajikan dalam Tabel 6. Dimana nilai *support*-nya menjadi informasi bagi pemilik toko *Groceries* dalam hal penempatan produk yang dijualnya. Sebagai contoh, produk *root vegetables, other vegetables* dan *whole milk* harus diletakkan berdampingan di etalase/rak-rak penjualan produk, karena paling sering dibeli secara bersamaan oleh konsumen, dengan peluang sebesar 0,02318. Produk *other vegetables, whole milk* dan *yogurt* harus diletakkan berdampingan di etalase/rak penjualan produk, karena berada di

Tabel 6. Nilai *Support* frequent 3-itemsets

No	Itemsets	Banyaknya Transaksi	<i>Support</i>
1	{root vegetables, other vegetables, whole milk}	228	0.02318
2	{other vegetables, whole milk, yogurt}	219	0.02227

Hasil Perhitungan *Confidence*

Pada bagian ini akan diuraikan perhitungan nilai *confidence*, yaitu suatu ukuran yang menunjukkan hubungan antara 2 *item* secara *conditional* (misal, menghitung kemungkinan seberapa sering *item* Y dibeli oleh pelanggan jika pelanggan tersebut membeli sebuah *item* X) di toko *Groceries*. Sebagai contoh untuk menghitung nilai *confidence* pada *itemsets* transaksi produk *other vegetables* dan *whole milk* yaitu 0.3867 Artinya peluang seseorang akan membeli produk *whole milk* jika sudah membeli produk *other vegetables* adalah 0.3867 atau sebesar 38,67%. 0.2928 Artinya peluang seseorang akan membeli produk *other vegetables* jika sudah membeli produk *whole milk* adalah 0.2928 atau sebesar 29,28%.

Selanjutnya untuk melihat nilai *confidence* frequent 3-itemsets yang dibeli oleh konsumen secara bersama-

Tabel 7. Nilai *Confidence* frequent 3-itemsets

Itemsets	Nilai <i>Confidence</i>					
	({BC} A)	({AC} B)	({AB} C)	(A {BC})	(B {AC})	(C {AB})
{root vegetables, other vegetables, whole milk}	0,2127	0,1198	0,0907	0,3098	0,474	0,4893
{other vegetables, whole milk, yogurt}	0,1151	0,0871	0,1596	0,3975	0,5129	0,2976

urutan ke-2 yang sering dibeli bersamaan oleh konsumen, dengan peluang sebesar 0,02227.

sama yang akan disajikan pada Tabel 7.

Sebagai contoh frequent 3-itemsets {root vegetables, other

vegetables dan whole milk} nilai confidence-nya dihitung menggunakan persamaan (2.2) yaitu 0.2127 Artinya peluang seseorang akan membeli produk

other vegetables dan whole milk

jika sudah membeli produk *root vegetables* adalah 0.2127 atau sebesar 21,27%. Berdasarkan Tabel 7 terlihat bahwa nilai confidence pada 3 ruas kolom sebelah kanan lebih besar dari 3 ruas kolom yang kiri, hal itu menunjukkan bahwa pihak manajer toko groceries harus menyodorkan produk dengan 2 itemset sekaligus, setelah itu barulah menawarkan produk/item lainnya sehingga konsumen yang ingin membeli produk tersebut mempunyai peluang yang lebih besar.

D. Kesimpulan

Dalam skripsi ini telah dibahas *market basket analysis* dengan algoritma *Eclat* yang diaplikasikan untuk data transaksi penjualan produk di toko *Groceries*. Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa:

1. Aplikasi algoritma *Eclat* dengan minimum support 2% dari keseluruhan transaksi didapatkan 59 frequent 1-Itemsets, 61 frequent 2-Itemsets dan 2 Frequent 3-Itemsets. Algoritma *Eclat* dihentikan di frequent 3-Itemsets karena sudah tidak ada lagi frequent Itemsets yang memenuhi minimum support 2% atau 197 transaksi.
2. Untuk frequent 2-Itemsets, produk *other vegetables* dan *whole milk* harus diletakkan berdampingan di etalase/rak-rak penjualan produk, karena paling sering dibeli secara bersamaan oleh konsumen, dengan peluang sebesar 0,07483.atau persentase jumlah transaksi sebesar 7,48%.

Untuk frequent 3-Itemsets, produk *root vegetables*,*other vegetables* dan *whole milk* harus diletakkan berdampingan di etalase/rak-rak penjualan produk, karena paling sering dibeli secara bersamaan oleh konsumen, dengan peluang sebesar 0,02318 atau persentase jumlah transaksi sebesar 2,32%.

3. Untuk frequent 2-Itemsets, alangkah lebih baik pihak manajer toko groceries menyodorkan produk *other vegetables* terlebih dahulu setelah itu barulah produk *whole milk*. Sedangkan untuk frequent 3-Itemsets, pihak manajer toko groceries lebih baik menyodorkan produk 2-Itemsets sekaligus setelah itu barulah menawarkan produk lainnya, yaitu *other vegetables* dan *whole milk* kemudian setelah itu barulah menawarkan produk *root vegetables*.

E. Saran

1. Disarankan kepada pemilik toko *Groceries* untuk menata produk yang dijual pada etalase/rak penjualan dimana menyesuaikan dengan hasil analisis yang peneliti lakukan yang berkaitan dengan produk-produk yang dibeli secara bersamaan.
2. Disarankan kepada peneliti lain untuk meneliti *Market Basket Analysis* dengan algoritma yang berbeda seperti algoritme *Hui Miner*, *Generalized Sequential Pattern* (GSP), dan lainnya.

Daftar Pustaka

- Albion Research Ltd. (2007). *Market Basket Analysis*. http://www.albionresearch.com/data_mining/market_basket.php. Diunduh tanggal 14 Juli 2018.
- Lantz, Brett. (2015). *Machine Learning*

With R real world grocery store.

- Larose , Daniel T, (2005), *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons. Inc.
- Olson, D., dan Shi, Y. (2008). *Pengantar Ilmu Penggalian Data Bisnis*. Salemba Empat, Jakarta
- Turban, E., dkk, (2005), *Decicion Support Systems and Intelligent Systems*.
- Zaki, M. J., Parthasarathy, S., Ogihara, M., & Li, W. (1997). New algorithms for fast discovery of association rules. In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining* (pp. 283–286). AAAI Press. Retrieved from <http://dl.acm.org/citation.cfm?id=3001392.3001454>