

# Pencarian *Frequent Itemset* pada Analisis Keranjang Belanja Menggunakan Algoritma FP-Growth

Lusa Indah Prahartiwi<sup>1,\*</sup>

<sup>1</sup>Sistem Informasi; STMIK Nusa Mandiri; Jl. Damai No. 8 Warung Jati Barat Jakarta Selatan DKI Jakarta, (telp: 021-78839502); e-mail: [lusaindah@gmail.com](mailto:lusaindah@gmail.com)

\* Korespondensi: e-mail: [lusaindah@gmail.com](mailto:lusaindah@gmail.com)

Diterima: 3 Oktober 2017 ; Review: 9 Oktober 2017 ; Disetujui: 13 November 2017

Cara sitasi: Prahartiwi LI. 2017. Pencarian *Frequent Itemset* pada Analisis Keranjang Belanja Menggunakan Algoritma FP-Growth. *Information System for Educators and Professionals*. 2 (1): 1 – 10.

**Abstrak:** Analisis keranjang belanja (juga dikenal sebagai *association rule mining*) adalah salah satu metode data mining yang berfokus untuk menemukan pola pembelian dengan mengekstrak asosiasi atau data transaksi dari sebuah toko. Analisis keranjang belanja menemukan produk yang dibeli bersamaan dalam keranjang yang sama. *Association rules* adalah suatu prosedur untuk mencari hubungan antar *item* yang ada pada suatu dataset. Penelitian ini menggunakan dataset Supermarket dan pengolahan data menggunakan perangkat lunak Rapid Miner. Metode yang digunakan dalam pencarian *frequent itemset* adalah Algoritma FP-Growth. Hasil eksperimen menggunakan Algoritma FP-Growth didapatkan bahwa kombinasi beer wine spirits-frozen foods dan snack foods merupakan *frequent itemset* dengan *lift ratio* sebesar 2.477.

**Kata kunci:** Analisis Keranjang Belanja, FP-Growth

**Abstract:** *Market basket analysis (also known as association rule mining) is one method of data mining that focuses on finding purchase patterns by extracting associations or transaction data from a store. Market basket analysis found products purchased together in the same bucket. Association rules is a procedure for finding relationships between items that exist on a dataset. This research uses Supermarket dataset and data processing using Rapid Miner software. The method used in the frequent itemset search is the FP-Growth Algorithm. Experimental results using FP-Growth Algorithm found that the combination of beer spirits-frozen foods and snack foods is a frequent itemset with an lift ratio of 2,477*

**Keywords:** *FP-Growth, Market Basekt Analysis*

## 1. Pendahuluan

Analisis keranjang belanja (juga dikenal sebagai *association rule mining*) adalah salah satu metode data mining yang berfokus untuk menemukan pola pembelian dengan mengekstrak asosiasi atau data transaksi dari sebuah toko [Kim et al., 2012]. Analisis keranjang belanja menemukan produk yang dibeli bersamaan dalam keranjang yang sama [Kim et al., 2012].

*Association Rules Mining* dapat menemukan *item* apa saja yang dibeli oleh pelanggan pada saat melakukan satu kali transaksi sehingga dapat mereorganisasi tata letak produk-produk yang saling berkaitan pada etalase yang berdekatan, dan juga untuk merancang promosi produk yang saling berhubungan. *Association Rules* menemukan semua aturan yang memiliki *support* dan *confidence* di atas batas yang ditentukan oleh pengguna [Chandra and Bhaskar, 2011].

Selama ini banyak supermarket yang dalam penataan produknya dilakukan dengan hanya mengelompokkan berdasarkan jenis produknya saja. Misal, produk susu ditempatkan pada satu etalase khusus susu dan berdekatan dengan etalase produk minuman lainnya. Padahal besar kemungkinan jika pelanggan datang membeli susu maka dia juga akan membeli

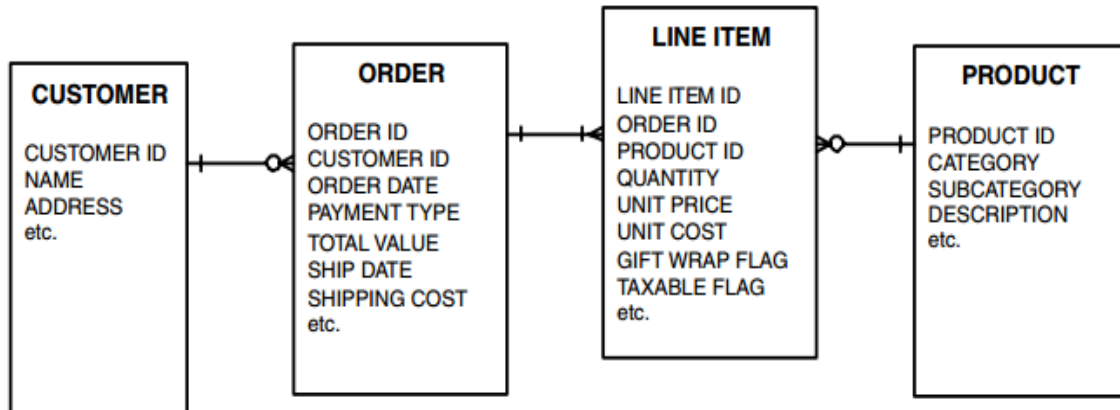
roti. Oleh karena itu diperlukan sebuah algoritma untuk menemukan pola berupa produk-produk yang dibeli secara bersamaan oleh pelanggan dalam satu kali transaksi. Penelitian ini bertujuan untuk mencari *frequent itemset* pada transaksi yang terjadi di Supermarket dengan menggunakan Algoritma FP-Growth.

### Data Mining

Data *mining* adalah proses menemukan korelasi baru, pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola serta statistic dan teknik matematika [Larose, 2006]. Data *mining* adalah analisis data pengamatan untuk menemukan hubungan tak terduga dan untuk meringkas data dalam cara-cara baru yang mudah dimengerti dan berguna untuk pemilik data [Hand et al., 2001]. Tujuan dari data *mining* adalah untuk menemukan pengetahuan yang tersembunyi dalam *dataset* dimana mata manusia atau analisis statistik konvensional tidak bisa melihatnya [Maione et al., 2016].

Peran Data *Mining* [Larose, 2006], diantaranya: Pertama, **Deskripsi**. Terkadang peneliti dan analis mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Kedua, **Estimasi**. Estimasi hampir sama dengan klasifikasi, kecuali variabel target numerik lebih baik dari pada kategori. Model yang dibangun menggunakan record lengkap yang memberikan nilai dari variabel target sebagai nilai prediksi. Kemudian, pada pengamatan selanjutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan usia pasien, jenis kelamin, berat badan, dan tingkat sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Selanjutnya model estimasi yang dihasilkan dapat diterapkan untuk kasus baru lainnya. Ketiga, **Prediksi**. Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa depan. Contoh prediksi dalam bisnis dan penelitian adalah prediksi harga saham dalam tiga bulan yang akan datang dan prediksi presentase kenaikan kematian akibat kecelakaan lalu lintas tahun depan jika batas kecepatan meningkat. Keempat, **Klasifikasi**. Dalam klasifikasi, terdapat target variabel kategori. Misalnya, penggolongan pendapatan dibedakan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah yang didasarkan pada karakteristik seperti usia, jenis kelamin dan pekerjaan. Contoh lain klasifikasi dalam bisnis dan penelitian adalah menentukan apakah suatu transaksi kartu kredit merupakan transaksi penipuan, memperkirakan apakah aplikasi hipotek merupakan risiko kredit yang baik atau buruk, mendiagnosa penyakit dan mengidentifikasi apakah ada perilaku tertentu yang menunjukkan kemungkinan ancaman teroris. Kelima, **Pengklusteran (Clustering)**. Pengklusteran merupakan pengelompokan catatan, pengamatan, atau kasus dari objek-objek yang memiliki kemiripan. Kluster adalah kumpulan catatan yang memiliki kemiripan satu sama lain dan berbeda dengan catatan dalam kelompok lainnya. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Sebaliknya, algoritma pengklusteran melakukan pengelompokan terhadap seluruh data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), dimana kemiripan catatan dalam kluster bernilai maksimal dan kemiripan catatan cluster lain akan bernilai minimal. Contoh pengklusteran dalam bisnis dan penelitian adalah Untuk tujuan audit akutansi, yaitu melakukan pemisahan terhadap perilaku finansial ke dalam kategori baik dan mencurigakan dan melakukan pengklusteran terhadap ekspresi dari gen, dalam jumlah yang sangat besar. Keenam, **Asosiasi**. Tugas asosiasi dalam data *mining* adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Masalah asosiasi pertama kali yang diusulkan adalah analisis keranjang belanja. Dalam masalah ini, diinginkan menentukan aturan terkait perilaku pembelian pelanggan [Aggarwal, 2015]. Setiap pelanggan membeli satu set yang berbeda dari produk, dalam jumlah yang berbeda, pada waktu yang berbeda Analisis keranjang belanja menggunakan informasi tentang apa yang pelanggan beli untuk memberikan pengetahuan siapa mereka dan mengapa mereka melakukan pembelian tertentu [Berry and Linoff, 2004]. Analisis keranjang belanja mengacu pada satu set masalah bisnis yang terkait dengan memahami *point-of-sale* data transaksi [Berry and Linoff, 2004]. Data keranjang belanja adalah data transaksi yang menggambarkan tiga entitas dasar yang berbeda, yaitu customer, orders dan items.

Dalam *database relational*, struktur data untuk data keranjang belanja sering terlihat seperti pada Gambar 1. Struktur data ini mencakup empat entitas penting, yaitu *Customer*, *Order*, *Line Item* dan *Product*.



Sumber: [Berry and Linoff, 2004]

**Gambar 1. Model Data Untuk Data Keranjang Belanja**

Teknik data mining yang paling dekat dengan masalah analisis keranjang belanja adalah *Association Rules Mining*

**Association Rules Mining**

*Association Rules* adalah suatu prosedur untuk mencari hubungan antar *item* yang ada pada suatu *dataset*. *Association Rules* berasal dari *point-of-sale* data yang menggambarkan produk apa yang dibeli bersama-sama [Berry and Linoff, 2004]. Salah satu daya tarik aturan asosiasi adalah kejelasan dan kegunaan hasil, yang dalam bentuk aturan tentang kelompok produk [Berry and Linoff, 2004]. Masalah yang paling penting dalam *Association Rule Mining* adalah penemuan aturan asosiasi dalam *database* transaksi [Chandra and Bhaskar, 2011]. Masalah klasik dalam *Association Rules Mining* adalah supermarket. Data yang berisi set *item* dibeli oleh pelanggan, yang disebut sebagai transaksi. Tujuannya adalah untuk menentukan hubungan antara kelompok barang yang dibeli oleh pelanggan, yang dapat secara intuitif akan melihat arah korelasi antara *item* [Aggarwal, 2015]. *Association rule* merupakan sebuah ekspresi implikasi yang berbentuk  $X \rightarrow Y$ , dimana X dan Y merupakan set dari *items* [Aggarwal, 2015]. X merupakan *antecedent* dan Y merupakan *consequent* [Gorunescu, 2011]. Perhitungan inti dalam *association rule mining* adalah untuk mengidentifikasi "*Frequent itemset*", yaitu, semua *itemset* yang *support* (frekuensi) dalam *database* adalah lebih dari batas *minimal support* yang ditentukan oleh pengguna [Aggarwal and Yu, 2008].

Proses *Association Rules Mining* terdiri dari 2 tahapan [Vu and Alaghband, 2014], yaitu: Pertama, Menambang semua pola *frequent* (*frequent itemset*), kedua menghasilkan aturan dari semua pola-pola tersebut. *Association Rules* mining membutuhkan dua hal penting yaitu *Minimum Support* dan *Minimum Confidence* [Bhandari et al., 2015]. *Support* digunakan untuk menentukan seberapa banyak aturan dapat diterapkan pada *dataset*, sedangkan *confidence* digunakan untuk menentukan seberapa sering *item* di dalam Y muncul dalam transaksi yang berisi X.

Rumus untuk mencari *support* dan *confidence* [Aggarwal, 2015];

$$Support (X \Rightarrow Y) = \frac{X \cup Y}{Jumlah\ transaksi\ dalam\ database} \dots\dots\dots(1)$$

$$Confidence (X \Rightarrow Y) = \frac{Sup(X \cup Y)}{Sup(X)} \dots\dots\dots(2)$$

### FP-Growth

FP-Growth merupakan metode pertambahan *frequent itemset* tanpa calon generasi. FP-Growth membangun struktur data (FP-Tree) untuk mengekstrak *database* transaksi [Han et al., 2012]. FP Growth melakukan *database* hanya dua kali. Di *scan* pertama, semua *item* yang *frequent* dan jumlah *support* berasal diurutkan dari *support* yang terbesar dalam setiap transaksi. Di *scan* kedua, *item* di setiap transaksi digabung menjadi FP-Tree dan *item* (node) yang muncul di transaksi yang berbeda dihitung. Setiap node berhubungan dengan *item* dan jumlahnya. Node dengan label yang sama dihubungkan oleh pointer yang disebut node-link [Wu and Kumar, 2009].

Tahapan dasar dalam metode FP-Growth terdiri dari dua tahap [Harrington, 2012], yaitu: *Conditional Pattern Base, Conditional FP-Tree, Frequent Pattern*.

Langkah-langkah metode FP-Growth: Pertama *Scan Database* pertama, kemudian hitung nilai *support* menggunakan rumus (1) untuk menentukan nilai *support* dari setiap *itemset*. *Itemset* yang tidak memenuhi syarat *minimum support* akan di hilangkan, kedua *Itemset* yang memperoleh nilai *support* lebih dari nilai *minimum support* di urutkan secara menurun, ketiga *scan Database* yang kedua untuk membentuk FP-Tree, keempat pembangkitan *Conditional Pattern Base* yang berisi *prefix path*(lintasan *prefix*) dan *suffix pattern*, kelima jumlahkan nilai *support* dari setiap *item* pada *conditional pattern base*, selanjutnya *item* yang memiliki nilai *support* yang lebih besar sama dengan dari *minimum support* akan di bangkitkan dengan *conditional fp-tree*. Apabila *conditional FP-Tree* merupakan lintasan tunggal, maka didapatkan *frequent itemset* dengan melakukan kombinasi *item* untuk setiap *conditional fp-tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-Growth secara rekursif

### Metode Evaluasi Association Rules Mining

Salah satu cara yang lebih baik untuk melihat kuat tidaknya aturan asosiasi adalah membandingkannya dengan nilai *benchmark*, dimana kita asumsikan kejadian *item* dari *consequent* dalam suatu transaksi adalah independen dengan kejadian dari *antecedent* dari suatu aturan asosiasi [Santosa, 2007].

$$\frac{P(\text{antecedent}) * P(\text{consequent})}{P(\text{antecedent})} = P(\text{consequent}) \dots\dots\dots(1)$$

*Confidence benchmark* dinyatakan dengan:

$$\frac{\text{jumlah transaksi dengan item dalam consequent}}{\text{jumlah transaksi dalam database}} \dots\dots\dots(2)$$

Membandingkan *confidence* terhadap *confidence benchmark* dengan melihat *rationya*, yang dinamakan *ratio lift*. Jadi *lift ratio* adalah perbandingan antara *confidence* untuk suatu aturan dibagi dengan *confidence*, dimana diasumsikan *consequent* dan *antecedent* saling independen [Santosa, 2007].

$$\text{Lift Ratio} = \frac{\text{confidence}}{\text{benchmark confidence}} \dots\dots\dots(3)$$

Jika *lift ratio* lebih besar dari 1, maka menunjukkan adanya manfaat dari aturan tersebut. Lebih tinggi nilai *ratio*, lebih besar kekuatannya [Santosa, 2007].

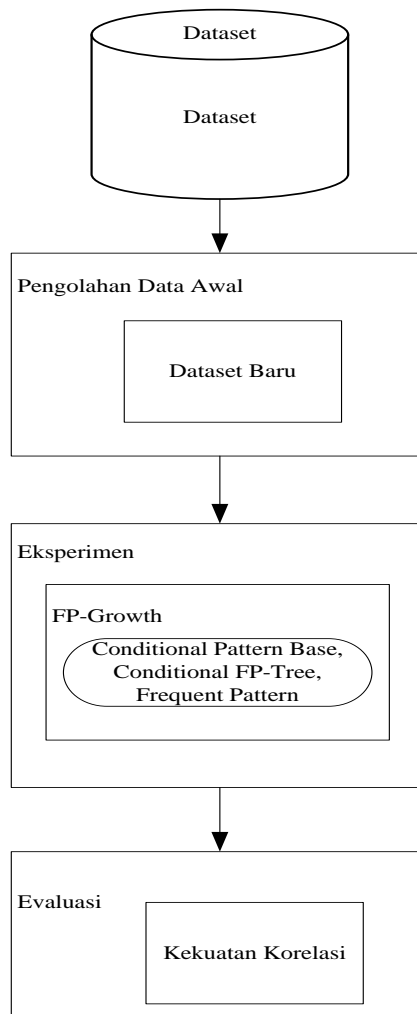
## 2. Metode Penelitian

Terdapat dua pendekatan utama dalam penelitian, yaitu pendekatan kualitatif dan pendekatan kuantitatif. Pendekatan kualitatif berhubungan dengan penilaian dari sikap, pendapat dan perilaku. Secara umum teknik yang digunakan adalah *interview* pada kelompok tertentu dan wawancara yang mendalam [Kothari, 2004]. Sedangkan pendekatan kuantitatif umumnya didorong oleh hipotesis, yang dirumuskan dan diuji, dengan tujuan menunjukkan bahwa hipotesis salah [Berndtsson et al., 2008].

Metode penelitian kuantitatif terbagi menjadi dua, yaitu metode Survei dan metode Eksperimen. Metode penelitian yang digunakan pada penelitian ini adalah metode

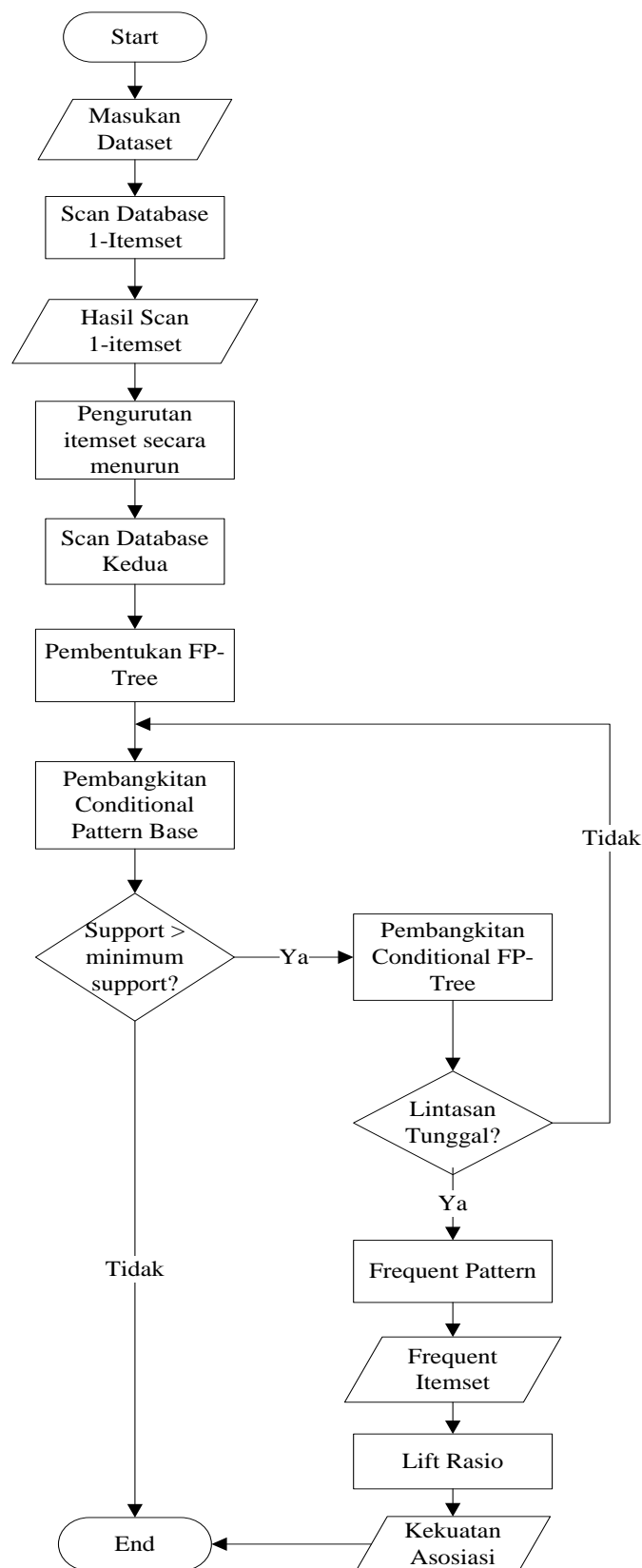
penelitian *Experimen*. Penelitian eksperimen melibatkan penyelidikan perlakuan parameter/variabel tergantung pada penelitiannya dan menggunakan tes yang dikendalikan oleh si peneliti itu sendiri

Tahapan-tahapan dalam penelitian ini dapat dilihat pada Gambar 2. Pertama Pengumpulan Dataset, data yang digunakan pada penelitian ini merupakan data sekunder. Data sekunder adalah data yang tidak diperoleh langsung dari obyek penelitian, melainkan telah dikumpulkan oleh pihak lain. Data sekunder yang digunakan dalam penelitian ini yaitu data transaksi Supermarket. Data transaksi supermarket ini sama dengan yang digunakan oleh [Wisaeng, 2014] dan [Hofman and Klinkenberg, 2009]. *Minimum support* ditetapkan sebesar 10% dan *minimum confidence* sebesar 70%. Kedua Pengolahan Data Awal, pengolahan data awal meliputi *Data Cleaning*, dan *Data Reduction*. Pada *dataset* supermarket satu atribut perlu dihapus yaitu atribut *Receipt Id* yang mana atribut tersebut dianggap tidak diperlukan setelah dilakukan seleksi fitur pada *dataset* supermarket. Ketiga Metode yang diusulkan, metode yang diusulkan adalah menerapkan Algoritma FP-Growth. Dataset akan diolah dengan metode yang diusulkan dan hasil pengukuran akan dianalisa dan dievaluasi. Keempat Eksperimen dan Pengujian Metode, penelitian yang dilakukan dalam eksperimen ini adalah dengan menggunakan komputer untuk melakukan proses perhitungan terhadap metode yang diusulkan. Kelima Evaluasi Hasil, metode yang diusulkan akan diuji dengan menggunakan *Lift ratio* untuk mengetahui kekuatan korelasinya. Kekuatan korelasi dihitung dari nilai *confidence* dibandingkan dengan *benchmark confidence*. Semakin tinggi nilai *lift ratio*, semakin baik pula kekuatan asosiasinya. Gambar 2 merupakan gambar alur metode yang diusulkan dalam penelitian ini. Sedangkan Algoritma dari metode yang diusulkan digambarkan dengan *flowchart* pada Gambar 3



Sumber: Hasil Penelitian (2017)

Gambar 2. Metode yang Diusulkan



Sumber: Hasil Penelitian (2017)

Gambar 3. Flowchart Metode yang Diusulkan

### 3. Hasil dan Pembahasan

Eksperimen algoritma FP-Growth pada dataset Supermarket dengan menggunakan tools Rapid Miner. *Minimum support* ditetapkan sebesar 10% dan *minimum confidence* ditetapkan sebesar 70%. Gambar 4 merupakan proses pengolahan dataset Supermarket untuk menghasilkan *frequent itemset* dan aturan asosiasi menggunakan Rapid Miner.



Sumber: Hasil Pengolahan Data (2017)

Gambar 4. Proses Eksperimen Algoritma FP-Growth pada Dataset Supermarket Menggunakan Tools Rapid Miner

Berdasarkan proses yang ditampilkan pada Gambar 3, diperoleh hasil aturan asosiasi dan *frequent itemsets*. Pada Gambar 5 dapat dilihat, bahwa yang menjadi *frequent itemset* adalah kombinasi *itemset* beer wine spirits-frozen foods dan snack foods. Nilai *support* yang dihasilkan sebesar 0.156 atau 15.6%, sedangkan nilai *confidence* sebesar 0.838 (83.8%). Kekuatan korelasi yang didapat oleh kombinasi beer wine spirits-frozen foods dan snack foods adalah sebesar 2.477. Karena hasil yang didapat lebih dari 1, maka menunjukkan adanya manfaat dari aturan asosiasi tersebut.

| Premises                                     | Conclusion        | Supp... ↓ | Confidence | LaPlace | Gain   | p-s   | Lift  |
|--|-------------------|-----------|------------|---------|--------|-------|-------|
| beer_wine_spirits, snack_foods               | frozen_foods      | 0.156     | 0.815      | 0.970   | -0.227 | 0.093 | 2.454 |
| snack_foods, frozen_foods                    | beer_wine_spirits | 0.156     | 0.825      | 0.972   | -0.222 | 0.092 | 2.431 |
| beer_wine_spirits, frozen_foods              | snack_foods       | 0.156     | 0.838      | 0.975   | -0.216 | 0.093 | 2.477 |
| beer_wine_spirits, snack_foods, meats        | frozen_foods      | 0.049     | 0.817      | 0.990   | -0.071 | 0.029 | 2.461 |
| beer_wine_spirits, frozen_foods, meats       | snack_foods       | 0.049     | 0.849      | 0.992   | -0.067 | 0.030 | 2.506 |
| snack_foods, produce, frozen_foods           | beer_wine_spirits | 0.044     | 0.793      | 0.989   | -0.067 | 0.025 | 2.338 |
| beer_wine_spirits, produce, frozen_foods     | snack_foods       | 0.044     | 0.799      | 0.989   | -0.067 | 0.026 | 2.359 |
| beer_wine_spirits, snack_foods, produce      | frozen_foods      | 0.044     | 0.825      | 0.991   | -0.063 | 0.026 | 2.486 |
| beer_wine_spirits, frozen_foods, paper_goods | snack_foods       | 0.040     | 0.765      | 0.988   | -0.065 | 0.022 | 2.259 |
| snack_foods, frozen_foods, paper_goods       | beer_wine_spirits | 0.040     | 0.787      | 0.990   | -0.062 | 0.023 | 2.317 |
| beer_wine_spirits, snack_foods, paper_goods  | frozen_foods      | 0.040     | 0.796      | 0.990   | -0.061 | 0.023 | 2.399 |
| beer_wine_spirits, snack_foods, breads       | frozen_foods      | 0.033     | 0.802      | 0.992   | -0.050 | 0.020 | 2.416 |
| beer_wine_spirits, frozen_foods, breads      | snack_foods       | 0.033     | 0.840      | 0.994   | -0.046 | 0.020 | 2.482 |
| snack_foods, frozen_foods, breads            | beer_wine_spirits | 0.033     | 0.853      | 0.994   | -0.045 | 0.020 | 2.514 |

Sumber: Hasil Pengolahan Data (2017)

Gambar 5. Hasil Eksperimen Algoritma FP-Growth pada Dataset Supermarket menggunakan Tools Rapid Miner



#### 4. Kesimpulan

Analisis keranjang belanja dapat dilakukan dengan pengetahuan *Association Rules Mining*. Eksperimen dilakukan dengan menggunakan Algoritma FP-Growth serta menggunakan dataset Supermarket. Dari hasil eksperimen diperoleh aturan asosiasi bahwa yang menjadi *frequent itemset* adalah kombinasi *itemset* beer wine spirits-frozen foods dan snack foods. Nilai *support* yang dihasilkan sebesar 0.156 atau 15.6%, sedangkan nilai *confidence* sebesar 0.838 (83.8%). Hal ini berarti, jika pelanggan membeli produk beer wine spirits dan frozen foods, maka pelanggan tersebut juga akan membeli produk snack foods. Sehingga pemilik supermarket dapat menempatkan produk-produk yang berkaitan tersebut saling berdekatan. Atau juga pemilik supermarket dapat melakukan promosi dengan pemberian diskon apabila pelanggan membeli beberapa produk beer wine spirit dan frozen foods maka akan mendapat bonus snack foods. Kekuatan korelasi yang didapat oleh kombinasi beer wine spirits-frozen foods dan snack foods adalah sebesar 2.477 yang menunjukkan adanya manfaat dari aturan asosiasi tersebut.

#### Referensi

- Aggarwal CC. 2015. Data Mining: The Textbook. 746 p.
- Aggarwal CC, Yu PS. 2008. Privacy-preserving data mining: models and algorithms. IEEE Secur. Priv. Mag. 2: 19–27.
- Berndtsson M, Hansson J, Olsson B, Lundell B. 2008. Thesis Projects: A Guide for Students in Computer Science and Information Systems. 165 p.
- Berry MJ a., Linoff GS. 2004. Data mining techniques: for marketing, sales, and customer relationship management. 545 p.
- Bhandari A, Gupta A, Das D. 2015. Improved apriori algorithm using frequent pattern tree for real time applications in data mining. In: Procedia Computer Science., p 644–651.
- Chandra B, Bhaskar S. 2011. A new approach for generating efficient sample from market basket data. Expert Syst. Appl. 38: 1321–1325.
- Gorunescu F. 2011. Data mining: Concepts, models and techniques. Intell. Syst. Ref. Libr. 12.
- Han J, Kamber M, Pei J. 2012. Data mining: concepts and techniques. 772 p.
- Hand DJ, Mannila H, Smyth P. 2001. Principles of Data Mining (Adaptive Computation and Machine Learning). J. Am. Stat. Assoc. 98: 252–253.
- Harrington P. 2012. Machine Learning in Action. 1-20 p.
- Hofman;M. 2009. Data Mining and Knowledge Discovery Series.
- Kim HK, Kim JK, Chen QY. 2012. A product network analysis for extending the market basket analysis. Expert Syst. Appl. 39: 7403–7410.
- Kothari CR. 2004. Research Methodology: Methods & Techniques. 418 p.
- Larose DT. 2006. Data Mining Methods and Models. 1-322 p.
- Maione C, De Paula ES, Gallimberti M, Batista BL, Campiglia AD, Barbosa F, Barbosa RM. 2016. Comparative study of data mining techniques for the authentication of organic grape juice based on ICP-MS analysis. Expert Syst. Appl. 49: 60–73.
- Vu L, Alagband G. 2014. Novel parallel method for association rule mining on multi-core shared memory systems. Parallel Comput. 40: 768–785.

Wisaeng K. 2014. Association rule with frequent pattern growth algorithm for frequent item sets mining. *Appl. Math. Sci.* 8: 4877–4885.

Wu X, Kumar V. 2009. *The Top Ten Algorithms in Data Mining*. 208 p.