

PENCARIAN POLA ASOSIASI DATA TRANSAKSI PENJULAN SPAREPART MOTOR MENGUNAKAN ALGORITMA FP-GROWTH

¹Wawan Adi Purwanto

²Agung Nilogiri

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

¹E-mail: adiwawan48@gmail.com

ABSTRAK

Pada saat ini dunia bisnis semakin berkembang pesat, terutama dalam dunia perdagangan, banyaknya data transaksi penjualan yang disimpan menyebabkan penumpukan data, yang mana data tersebut dapat proses lebih lanjut menjadi suatu informasi yang berguna dalam pengambilan keputusan. PT. Part Station merupakan perusahaan di Indonesia yang bergerak dibidang *retail* dengan produk utama otomotif. Perusahaan tersebut memiliki kendala dalam mengelola penempatan barang yang beragam, dari banyaknya barang yang terdapat pada data transaksi penjualan, dapat diketahui kebiasaan pelanggan mengenai barang apa saja yang sering dibeli secara bersamaan. Oleh sebab itu perusahaan harus mengambil keputusan yang tepat dalam menentukan strategi pengaturan tata letak barang untuk meletakkan barang secara optimal, agar karyawan bisa melayani pembeli dengan lebih cepat dan mudah. Dengan adanya metode *association rule* diharapkan dapat membantu perusahaan mengetahui berapa *rule* yang dihasilkan dengan batasan *minimum support* dan *minimum confidence* untuk rekomendasi peletakan barang. Algoritma *FP-Growth* merupakan algoritma yang sangat efisien dalam pencarian *frequent itemset* dalam sebuah kumpulan data dengan membangkitkan struktur *prefix-tree* atau disebut dengan *FP-Tree*. Hasil perhitungan dari 383 data transaksi tahun 2017 dengan eksperimen *minimum support* 10% sampai dengan 100% dan *minimum confidence* 10% sampai dengan 100% dengan kenaikan peringkat 10% Ditemukan 4 *rule* pada *minimum support* 10% dan *minimum confidence* 10% sampai dengan 20%, sedangkan 2 *rule* ditemukan pada *minimum confidence* 30% sampai dengan 60%, sedangkan 1 *rule* ditemukan pada *minimum confidence* 70%. Sedangkan *rule* tidak ditemukan pada *minimum confidence* 80% sampai dengan 100% dan Pada *minimum support* 20% atau lebih.

Kata Kunci: tata letak barang, *association rule*, algoritma *FP-Growth*.

ABSTRACT

At this time the business world is growing rapidly, especially in commerce business, a large amount of sales transaction data stored causes accumulation of data, which data can further process into useful information in decision making. PT. Part Station is a company in Indonesia engaged in retail with major automotive products. The company has obstacles in managing the placement of various items, from the number of items contained in the sales transaction data, it can be seen the customer's habits regarding what items are often purchased simultaneously. Therefore the company must make the right decision in determining the strategy of arranging the layout of the goods to place the items optimally, so that employees can serve buyers more quickly and easily. With the existence of the association rule method, it is expected to help companies find out how many rules are produced with a minimum support limit and minimum confidence for laying recommendations. FP-Growth Algorithm is a very efficient algorithm in searching frequent itemset in a data set by generating prefix-tree structure or called FP-Tree. Calculation results from 383 transaction data for 2017 with minimum support 10% up to 100% experimentation and 10% up to 100% minimum confidence with 10% increase Found 4 rules at 10% minimum support and 10% minimum confidence up to 20%, while 2 rules are found with a minimum confidence of 30% to 60%, while 1 rule is found with a minimum confidence of 70%. While the rule is not found at a minimum confidence of 80% to 100% and at a minimum support of 20% and above.

Keywords: *item layout, association rule, FP-Growth algorithm.*

1. Pendahuluan

Pada saat ini dunia bisnis semakin berkembang pesat, terutama dalam dunia perdagangan. Sehingga para pengembang dan pelaku bisnis harus mencari solusi dan memikirkan strategi dan terobosan yang dapat menjamin keberlangsungan bisnis perusahaan. Salah satu cara yang bisa dilakukan untuk mengetahui kondisi pasar (pelanggan) adalah dengan mengamati data transaksi penjualan. Data transaksi penjualan disimpan dalam basis data server dan kemudian data inilah yang diproses sehingga dihasilkan informasi baru (Erwin, 2009).

PT. Part Station merupakan salah satu perusahaan bengkel terbesar yang ada di Indonesia, yang bergerak dibidang *retail* dengan produk utama otomotif. Perusahaan tersebut memiliki kendala dalam mengelola penempatan barang yang beragam, oleh sebab itu perusahaan harus mengambil keputusan yang tepat dalam menentukan strategi pengaturan tata letak barang untuk meletakkan item barang secara optimal, agar karyawan bisa melayani pembeli dengan lebih cepat dan mudah (Amrin, 2017). Untuk dapat melakukan hal tersebut, perusahaan membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut. Dengan memanfaatkan data transaksi penjualan yang telah tersimpan dalam *database*, pihak manajemen dapat mengetahui kebiasaan pelanggan atau perilaku pelanggan mengenai apa saja barang otomotif yang sering dibeli secara bersamaan.

Dengan menggunakan data mining, data tersebut dapat lebih dioptimalkan pemanfaatannya dengan mencari informasi yang tersembunyi dan jarang diketahui. Informasi tersebut dapat digunakan untuk meningkatkan keunggulan dalam persaingan bisnis retail. Salah satu fungsi data mining adalah *association rule*, yaitu fungsi untuk mencari informasi berupa asosiasi atau hubungan antar item dalam suatu data transaksi dan menampilkannya dalam bentuk pola beli konsumen dalam berbelanja. Pengetahuan mengenai pola inilah yang

nantinya bisa menjadi pedoman untuk meningkatkan keunggulan dalam persaingan bisnis *retail*, dengan cara mengoptimalkan tata letak barang yang sesuai dengan pola beli konsumen, sehingga dapat memudahkan karyawan dalam melayani pembeli. Suatu pola ditentukan oleh dua parameter, yaitu *Support* dan *Confidence*.

1.1 Data Mining

Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* besar (Larose, 2005).

Data mining sering disebut juga *Knowledge Discovery in Database* atau disingkat dengan KDD, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam dataset berukuran besar (Santosa, 2007).

1.2 Association Rule

Association rule adalah metode untuk mencari hubungan antar item yang terdapat dalam data yang besar (Aprilia, et al, 2013). Dalam pencarian pola asosiasi dari suatu kumpulan data, tahap pertama yang dilakukan adalah mengetahui jumlah *item* dari keseluruhan transaksi, setelah itu mencari *support* dan *confidence*.

1) Support

Support adalah nilai penunjang dari kombinasi *item* tersebut, dimana jika mempunyai item A dan item B maka *support* adalah nilai penunjang dari transaksi dalam data yang mengandung A dan B. Rumus untuk menghitung nilai *support* 1 *item* adalah sebagai berikut (Kusrini, 2009):

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \times 100\%$$

Sementara itu, nilai *support* dari 2 *item* diperoleh dari rumus berikut:

$$\text{Support (A} \cap \text{B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}} \times 100\%$$

2) Confidence

Confidence adalah ukuran ketepatan suatu *rule*, yaitu presentasi transaksi dalam data yang mengandung A dan mengandung B. Dengan adanya *confidence* kita dapat mengukur kuatnya hubungan antar *item* dalam aturan asosiasi. Rumus untuk menghitung nilai *confidence* adalah sebagai berikut (Kusrini, 2009):

$$\text{Confidence } P(A|B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi mengandung B}}$$

1.3 Algoritma Frequent Pattern Growth

Algoritma *FP-Growth* merupakan pengembangan dari algoritma Apriori. Algoritma *Frequent Pattern Growth (FP-Growth)* adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. (Samuel, 2008).

Pada algoritma *FP-Growth* menggunakan konsep pembangunan *tree*, yang biasa disebut *FP-Tree*, dalam pencarian *frequent itemset* bukan menggunakan *generate candidate* seperti yang dilakukan pada algoritma Apriori. Dengan menggunakan konsep tersebut, algoritma *FP-Growth* menjadi lebih cepat daripada algoritma Apriori. (Erwin, 2009).

Algoritma *FP-Growth* memiliki tahapan-tahapan yang harus dilewati agar dapat memberikan hasil yang maksimal, tahapan-tahapan tersebut (Han, 2006), yaitu:

- 1) Tahap pembangkitan *conditional pattern base*.
- 2) Tahap pembangkitan *conditional FP-Tree*.
- 3) Tahap pencarian *frequent itemset*.

1.4 Lift Ratio

Lift ratio merupakan ukuran yang digunakan untuk mengukur tingkat keakuratan sebuah *rule* yang dihasilkan dari proses asosiasi. Rumus yang digunakan untuk menghitung *lift ratio* adalah sebagai berikut (Makani et al, 2013):

$$\text{Lift Ratio } (A \rightarrow B) = \frac{\text{Support } (A \cap B)}{\text{Support } (A) \times \text{Support } (B)}$$

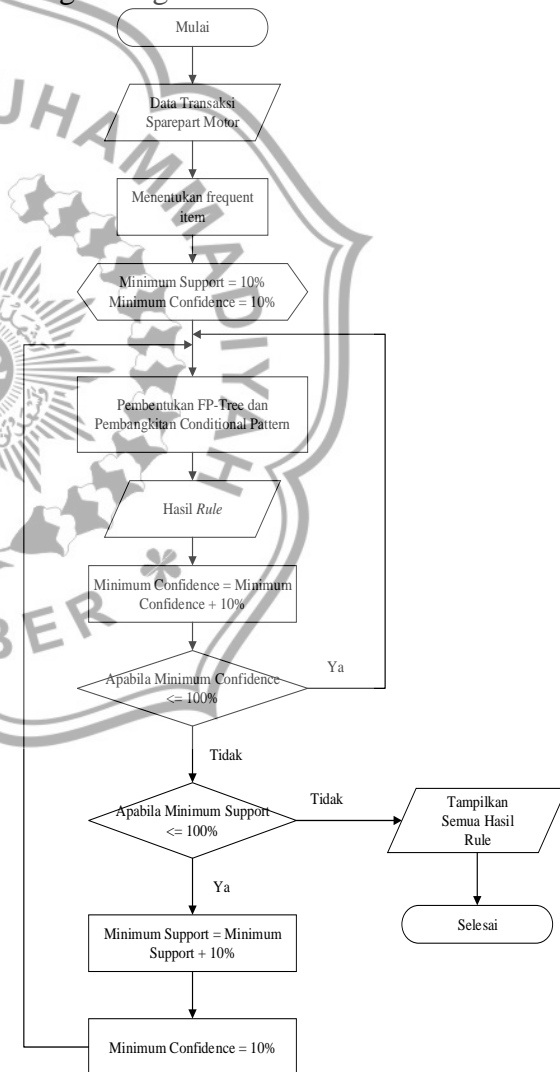
2. Metode Penelitian

2.1 Dataset

Pengumpulan data transaksi penjualan *sparepart motor* tahun 2017 di PT. Part Station cabang Jember diperoleh sebanyak 1798 data transaksi, tetapi hanya 383 transaksi yang dipakai untuk perhitungan dengan algoritma *FP-Growth*. Karena pada tahap *preprocessing data*, nota yang berisi *item* tunggal akan dihapus.

2.2 Diagram Flowchart

Untuk menentukan aturan asosiasi dengan algoritma *FP-Growth* pada data transaksi penjualan *sparepart motor*, dapat ditunjukkan dengan diagram berikut:



Langkah-langkah dalam melakukan eksperimen parameter *minimum support* dan *minimum confidence* dalam pencarian aturan asosiasi dengan algoritma *FP-Growth*, dimana

yang pertama kali dilakukan adalah input dataset dan menentukan *frequent item* dari dataset, selanjutnya yaitu memberi nilai parameter *minimum support* dan *minimum confidence* dengan kelipatan 10%, setelah itu pembentukan *FP-Tree* dan melakukan pembangkitan *frequent pattern* diantaranya adalah pembangkitan *conditional pattern base*, pembangkitan *conditional FP-Tree*, *frequent itemset*, tahap-tahap selanjutnya adalah proses pencarian aturan asosiasi dengan eksperimen *minimum support* 10% sampai dengan 100% dan *minimum confidence* 10% sampai dengan 100%, dan langkah terakhir yaitu menampilkan semua aturan asosiasi yang di dapat dari eksperimen tersebut dengan menggunakan algoritma *FP-Growth*.

2.3 Proses Algoritma *FP-growth*

1) Pemodelan data

Sebelum dataset dihitung dengan algoritma *FP-Growth* maka dilakukan Pemodelan data transaksi penjualan *sparepart* motor, kemudian dikonversikan ke dalam bentuk 0 dan 1 atau bentuk biner. Dimana 0 adalah *sparepart* motor yang tidak dibeli dan 1 adalah *sparepart* motor yang dibeli.

2) Menentukan *Frequent item*

Frequent item ditentukan dengan menghitung jumlah dari setiap barang dari keseluruhan transaksi.

3) Menentukan *minimum support*

jika hasil perhitungan *support* 1 *item* tidak memenuhi dari *minimum support* yang telah ditentukan, maka *item* tersebut akan dihapus atau dibuang. Sehingga didapatkan *frequent item* yang memenuhi *minimum support* dan diurutkan berdasarkan *frequent* terbesar ke *frequent* terkecil.

4) Membangun *FP-Tree*

Tahap ini adalah tahap dimana setiap *item* yang memenuhi *minimum support* yang telah ditentukan, kemudian dibangun menjadi sebuah pohon *tree*. Pembangunan pohon *tree* berdasarkan dari setiap transaksi, dimana pembentukannya dimulai dari *frequent* terbesar ke *frequent* terkecil dalam suatu transaksi untuk diletakkan dibawah *root*, untuk *frequent* terbesar

akan menjadi *node* induk, dilanjutkan pada *item* selanjutnya sehingga terbentuk lintasan transaksi pada *tree*.

5) Pembangkitan *conditional pattern base*

Pembangkitan *conditional pattern base* didapatkan melalui pembentukan *FP-Tree* yang telah dibangun sebelumnya. Pengambilan *conditional pattern base* adalah melihat jumlah *frequent item* terkecil terlebih dahulu, kemudian *node* di atasnya *frequent item* terkecil tersebut pada satu lintasan pohon *tree* dibangkitkan.

6) Pembangkitan *conditional FP-Tree*

Pada tahap ini, *frequent* dari setiap *item* yang sama pada *conditional pattern base* dijumlahkan, jika *item* tunggal maka akan langsung dibangkitkan pada *conditional FP-Tree*.

7) Pencarian *frequent itemset*

Apabila *conditional FP-Tree* merupakan lintasan tunggal, maka didapatkan pencarian *frequent itemset* dengan melakukan kombinasi *item* untuk setiap *conditional FP-Tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan *FP-Growth* secara rekursif (proses memanggil dirinya sendiri).

8) Menghitung nilai *support*

Menghitung nilai *support* berdasarkan dari hasil pencarian *frequent item* yang didapat. Untuk menghitung nilai *support* dengan rumus sebagai berikut:

$$\text{Support } (A \cap B) = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Total Transaksi}} \times 100\%$$

9) Menghitung nilai *confidence*

Menghitung nilai *confidence* berdasarkan dari hasil pencarian *frequent item* yang didapat. Untuk menghitung nilai *confidence* dengan rumus sebagai berikut:

$$\text{Confidence } P(A|B) = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Total Transaksi mengandung } B}$$

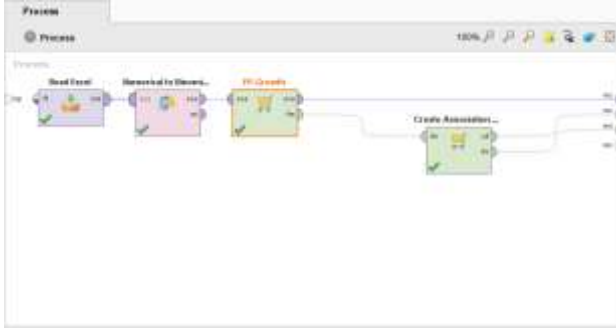
10) Menghitung nilai *lift ratio*

Menghitung nilai *Lift ratio* berdasarkan dari hasil pencarian *frequent item* yang didapat. Untuk menghitung nilai *Lift ratio* dengan rumus sebagai berikut:

$$\text{Lift Ratio } (A \rightarrow B) = \frac{\text{Support } (A \cap B)}{\text{Support } (A) \times \text{Support } (B)}$$

3. Hasil dan Pembahasan

3.1. RapidMiner



Gambar diatas adalah model algoritma FP-Growth dalam pencarian aturan asosiasi dengan menggunakan RapidMiner.

3.2. Hasil Aturan Asosiasi

Pengujian yang dilakukan dari 383 data transaksi dengan menggunakan bantuan aplikasi *RapidMiner Studio 9.4*. Parameter yang digunakan dengan *minimum support* 10% sampai dengan 100% dan *minimum confidence* 10% dengan sampai dengan 100% dengan kenaikan peringkat 10% didapatkan sebuah aturan asosiasi seperti pada tabel dibawah ini.

| Barang | Barang | Support | Confidence | Lift Ratio |
|--------|--------|---------|------------|------------|
| Bohlam | Oli | 11,5% | 20,6% | 1,291 |
| Busi | Oli | 16,2% | 29% | 1,156 |
| Oli | Busi | 16,2% | 63,9% | 1,144 |
| Oli | Bohlam | 11,5% | 72,1% | 1,291 |

Tabel diatas adalah hasil Pencarian aturan asosiasi dengan parameter *minimum support* 10% dan *minimum confidence* 10% sampai dengan 20%.

| Barang | Barang | Support | Confidence | Lift Ratio |
|--------|--------|---------|------------|------------|
| Oli | Busi | 16,2% | 63,9% | 1,144 |
| Oli | Bohlam | 11,5% | 72,1% | 1,291 |

Tabel diatas adalah hasil Pencarian aturan asosiasi dengan parameter *minimum support* 10% dan *minimum confidence* 30% sampai dengan 60%.

| Barang | Barang | Support | Confidence | Lift Ratio |
|--------|--------|---------|------------|------------|
| Oli | Bohlam | 11,5% | 72,1% | 1,291 |

Tabel diatas adalah hasil Pencarian aturan asosiasi dengan parameter *minimum support* 10% dan *minimum confidence* 70%.

Pencarian aturan asosiasi tidak ditemukan jika menggunakan *minimum support* 10% dan *minimum confidence* 80% sampai dengan 100%, sedangkan untuk *minimum support* 20% sampai dengan 100% dan *minimum confidence* 10% sampai dengan 100% tidak ditemukan hasil aturan asosiasi yang memenuhi parameter *minimum support* dan *minimum confidence*.

Hasil *lift ratio* sebelum implementasi untuk semua aturan asosiasi yang memenuhi parameter *minimum support* dan *minimum confidence* memiliki nilai *lift ratio* lebih besar dari 1. Hal ini menunjukkan bahwa aturan asosiasi yang memenuhi parameter *minimum support* dan *minimum confidence* bersifat kuat untuk digunakan sebagai acuan dalam mengatur tata letak barang pada rak barang.

4. Kesimpulan dan Saran

Berdasarkan hasil yang diperoleh, didapat kesimpulan sebagai berikut:

- 1) Untuk keluaran dengan 4 rule didapat pada *minimum support* 10% dan *minimum confidence* 10% sampai dengan 20%.
- 2) Untuk keluaran dengan 2 rule didapat pada *minimum support* 10% dan *minimum confidence* 30% sampai dengan 60%.
- 3) Untuk keluaran dengan 1 rule didapat pada *minimum support* 10% dan *minimum confidence* 70%.
- 4) Rule tidak ditemukan jika menggunakan *minimum support* 10% dengan *minimum confidence* 80% sampai dengan 100% dan *minimum support* 20% sampai dengan 100% dengan *minimum confidence* 10% sampai dengan 100%.
- 5) Berdasarkan hasil aturan asosiasi yang memenuhi *minimum support* 10% dan *minimum confidence* 10% sampai dengan 70% dan memiliki nilai *lift ratio* lebih dari 1 adalah sebanyak 4 aturan asosiasi, hal ini menunjukkan bahwa semua aturan asosiasi tersebut bersifat kuat untuk digunakan.

Saran dari penulis untuk penelitian selanjutnya yaitu:

- 1) Menggunakan metode asosiasi dengan algoritma yang berbeda seperti *algoritma Eclat*, *algoritma Fast Eclat*, *algoritma Fold Growth*, sehingga dapat dicari algoritma yang paling efektif dan efisien.
- 2) Dalam satu transaksi usahakan lebih dari satu *item* dengan frekuensi berulang lebih banyak dalam setiap transaksi.

5. Daftar Pustaka

- Amrin. (2017). *Data Mining dengan Algoritma Apriori untuk Penentuan Aturan Asosiasi Pola Pembelian Pupuk*. Amik Bina Sarana Informatika. Jakarta.
- Aprilia, D., Baskoro, D. A., Ambarwati, L., Wicaksana, I. S. (2013). *Belajar Data Mining Dengan Rapid Miner*. Gramedia Pustaka Utama. Jakarta
- Erwin. (2009). *Analisis Market Basket dengan Algoritma Apriori dan FP-Growth*. Jurnal Generic, 4.
- Han, J., Kamber, M. (2006). *Data Mining: Concepts and Techniques Second Edition*. San Fransisc: Morgan Kaufmann Publishers.
- Kusrini., Emha, L. T. (2009). *Algoritma Data Mining*. Penerbit Andi. Yogyakarta.
- Larose, D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. New Jersey: John Willey&Sons. Inc.
- Makani, Z., Arora, S., Kanikar, P. (2013). *A Parallel Approach to Combined Association Rule Mining*. International Journal of Computer Applications 7-13.
- Samuel, D. (2008). *Penerapan Struktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung. Bandung.
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Graha Ilmu. Yogyakarta.