

# PENERAPAN METODE ASSOCIATION RULE DENGAN ALGORITMA FP-GROWTH UNTUK MENGETAHUI KARAKTERISTIK PEMBELI MOBIL

Kadek Andriana Putra<sup>1</sup>, Drs.Mardji<sup>2</sup>, Lailil Muflikhah<sup>3</sup>

Jurusan Ilmu Komputer Program Teknologi Informasi dan Ilmu Komputer

Universitas Brawijaya

Jalan MT. Haryono 167, Malang 65145, Indonesia

E-mail: adeq.vhayer@gmail.com

## ABSTRAK

Persaingan didalam bisnis membutuhkan suatu strategi agar dapat meningkatkan penjualan. Strategi bisnis dapat dilakukan dengan segmentasi pasar berdasarkan aspek demografis dengan menganalisa karakteristik pembeli mobil berdasarkan parameter jenis kelamin dan rentang umur pada data transaksi penjualan sehingga diperoleh informasi sesuai dengan permintaan pasar. Penggalian data cukup efektif untuk mengetahui adanya serangkaian pola informasi dari data dalam jumlah besar. Dalam penelitian ini, penggalian dilakukan terhadap data transaksi penjualan *dealer* mobil. Sistem akan melakukan pencarian aturan asosiasi dengan tahap awal adalah mencari frequent itemset dengan menerapkan algoritma *fp-growth*. Kemudian dari *frequent itemset* tersebut akan dilakukan pencarian pola asosiasi dengan menggunakan teknik *association rule*. Pada tahap akhir, sistem akan menghitung kekuatan pola asosiasi yang terbentuk dengan menggunakan *lift ratio*. Pada penelitian ini didapatkan hasil pengujian nilai kekuatan (*lift ratio*) *rule*, jumlah *rule* yang dihasilkan tiap *dealer* adalah *dealer* Suzuki 3 *rule*, *dealer* Honda 5 *rule*, *dealer* Toyota 4 *rule*, dan *dealer* Daihatsu 4 *rule*. Dan didapatkan nilai rata – rata *lift ratio* dari *rule* yang dihasilkan dengan minimum confidence 50% sebesar 1.15. Nilai *lift ratio* tertinggi sebesar 2.19 pada *dealer* Honda dan nilai *lift ratio* terendah sebesar 0.90 pada *dealer* Toyota.

Kata kunci: penggalian data, aturan asosiasi, *fp-growth*, *lift ratio*.

## ABSTRACT

*Competition in business requires a strategy in order to increase sales. Business strategy can be done by segmenting the market based on demographic aspects by analyzing the characteristics of car buyers based on the parameters gender and age range of the sales transaction data to obtain information which fit with market demand. . Data mining is quite effective for the presence of a series of pattern information from large amounts of data. In this research, mining of data carried out on the car dealer sales transaction data. The data, performed to obtain the relationship between gender and age range of the type of car purchased. The system will do a search with the early stages of association rules is to find frequent itemset by applying fp-growth algorithm. Then association pattern searches will be conducted by using association rule after frequent itemset found. In the final stage, the system will calculate the power of association patterns with lift ratio equation. In this research, the results of testing the value of force (lift ratio) rule, the number of rules generated by the dealer is the dealer of Suzuki is 3 rules, dealer of Honda is 5 rules, dealer of Toyota is 4 rules, and dealer of Daihatsu is 4 rules. And average value earned on lift ratio of generated rule is 50% and minimum confidence of 1.15. Highest lift value ratio of 2.19 at dealer of Honda and the lowest value ratio of 0.90 at dealer of Toyota.*

*Keywords: data mining, association rules, fp-growth, lift ratio.*

## I. PENDAHULUAN

Banyaknya *dealer* mobil saat ini, membuat tiap *dealer* mobil harus berupaya untuk bersaing dalam meningkatkan penjualan. Untuk tetap dapat bersaing dalam pangsa pasar mobil teratas maka dibutuhkan informasi untuk mendukung pembuatan

strategi pemasaran dimana salah satunya adalah dengan melakukan segmentasi pasar. Segmentasi pasar menurut Kotler & Armstrong (2004) adalah membagi sebuah pasar ke dalam kelompok-kelompok pembeli yang khas berdasarkan kebutuhan, karakteristik, atau perilaku yang mungkin membutuhkan produk atau bauran pemasaran

yang terpisah. Melalui analisa karakteristik pembeli mobil, suatu *dealer* mobil memiliki tiap jenis mobil sehingga *dealer* mobil dapat mempertahankan bahkan meningkatkan tingkat penjualan serta memberikan dampak jangka panjang yang lebih kompetitif. Data mining merupakan salah satu proses analisa yang paling efektif untuk mengetahui pola informasi segmen-segmen pelanggan potensial dengan kesamaan karakteristik customer mobil dari sejumlah besar data yang ada. Dalam *data mining* terdapat beberapa teknik dimana penggunaannya bergantung terhadap permasalahan yang dihadapi. Pada permasalahan segmentasi pasar untuk mengetahui karakteristik pembeli mobil ini digunakan teknik *association rule*. Teknik *association rule* adalah teknik *data mining* yang dapat mencari korelasi antara *item* yang satu dengan *item* yang lain dengan mengetahui pola asosiasinya. Teknik *association rule* terdiri dari dua tahap, antara lain tahap pencarian kombinasi *item* yang paling sering muncul (pencarian *frequent itemset*) dan tahap pembangkitan *rule* yang telah terbentuk dari *frequent itemset* (Han dan Kamber, 2001). Salah satu teknik yang terdapat pada *association rule* adalah *FP-Growth*. Pencarian *frequent itemset* dengan algoritma *FP-Growth* tidak dengan membangkitkan kandidat melainkan dengan menggunakan konsep pembangunan *tree* (Erwin, 2009). Untuk dapat membangun *tree*, algoritma *FP-Growth* mencari *item* yang *frequent* (*frequent 1-itemset*) dengan cara membandingkan frekuensi kemunculannya dengan nilai minimum *support* yang telah ditentukan. *Item* yang *frequent* selanjutnya akan dibentuk kombinasinya sesuai dengan data transaksi yang ada sehingga untuk *item* yang tidak *frequent* telah dipangkas terlebih dahulu. Hasil kombinasi tersebut digunakan untuk membangun *tree* dan dianggap *frequent itemset* jika frekuensi kemunculannya lebih besar atau sama dengan nilai minimum *support* yang telah ditentukan. Data input yang dibutuhkan oleh algoritma *FP-Growth* harus memiliki *itemset* dengan kejadian yang berulang – ulang.

## II. TINJAUAN PUSTAKA

### 2.1 Association Rule

*Association Rule* adalah teknik data mining untuk menemukan aturan hubungan asosiatif antara suatu kombinasi *item* dengan

informasi tentang target pembeli mobil untuk

tujuan akhir mendapatkan kombinasi *item* yang dikatakan menarik (*interesting association rules*), terdiri dari :

#### a. Support

*Support* adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *itemset* dari keseluruhan transaksi (Han dan Kamber, 2001). Menurut Rochmah (2010), *support* dari aturan asosiasi adalah proporsi transaksi yang mengandung *antecedent* dan *consequent*. *Support* untuk aturan “X → Y” adalah probabilitas atribut atau kumpulan atribut X dan Y yang terjadi bersamaan dalam suatu transaksi (Yulita, 2002). Nilai *support* untuk suatu *item* dapat diperoleh dengan rumus pada Persamaan 2.1, sedangkan nilai *support* untuk kombinasi *item* *antecedent* dan *consequent* dapat diperoleh dengan rumus pada Persamaan 2.2 (Rochmah, 2010).

$$\text{Support}(A) = \frac{\text{Jumlah transaksi yang mengandung } A}{\text{Total transaksi}} \quad (2.1)$$

$$\text{Support}(A, C) = \frac{P(A \cap C)}{\text{Total transaksi}} = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } C}{\text{Total transaksi}} \quad (2.2)$$

Nilai *support* digunakan untuk menentukan *frequent itemset*. *Itemset* yang nilai *support*-nya memenuhi parameter minimum *support* (*min\_sup*) masuk dalam *frequent itemset*. Minimum *support* merupakan parameter yang digunakan sebagai batasan frekuensi kejadian atau jumlah *support* yang harus dipenuhi suatu kelompok data yang dijadikan aturan (Yulita, 2002).

#### b. Confidence

*Confidence* atau tingkat kepercayaan merupakan probabilitas kejadian beberapa produk yang dibeli secara bersamaan dimana salah satu produk sudah pasti dibeli, misalnya jika terdapat n transaksi dimana X yang dibeli dan m transaksi dimana X dan Y yang dibeli bersamaan, maka *confidence* dari aturan X → Y adalah m/n (Yulita, 2002). Nilai *confidence* dapat diperoleh dengan rumus pada Persamaan 2.3 dan dijabarkan lebih khusus pada Persamaan 2.4 (Rochmah, 2010).

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support } A \cap B}{\text{Support } A} \quad (2.3)$$

$$\text{Confidence}(A \rightarrow B) = P(B|A) = \frac{\text{Jumlah transaksi yang mengandung } A \cap B}{\text{Jumlah transaksi yang mengandung } A} \quad (2.4)$$

Nilai *confidence* digunakan dalam menentukan *strong association rule*. *Association rule* yang nilai *confidence*-nya memenuhi parameter *threshold* minimum *confidence* (*min\_conf*) termasuk dalam *strong association rule*. Minimum *confidence* adalah parameter yang mendefinisikan minimum *level* dari *confidence* yang dipenuhi oleh aturan yang berkualitas (Yulita, 2002).

**c. Lift Ratio**

*Lift ratio* digunakan untuk mengevaluasi kuat tidaknya sebuah aturan asosiasi. *Lift ratio* adalah perbandingan antara *confidence* sebuah aturan dengan nilai *benchmark confidence*. *Benchmark confidence* adalah perbandingan antara jumlah semua *item* yang menjadi *consequent* terhadap total jumlah transaksi (Santosa, 2007). Rumus *benchmark confidence* dapat dilihat pada Persamaan 2.5, sedangkan rumus *lift ratio* dapat dilihat pada Persamaan 2.6.

$$Benchmark\ Confidence = \frac{Nc}{N} \tag{2.5}$$

Keterangan rumus 2.5 :

Nc = jumlah transaksi dengan *item* yang menjadi *consequent*

N = jumlah transaksi basis data

$$Lift\ Ratio = \frac{Confidence(A,C)}{Benchmark\ Confidence(A,C)} \tag{2.6}$$

Apabila nilai *lift ratio* lebih besar dari 1, maka menunjukkan adanya manfaat dari aturan tersebut. Lebih tinggi nilai *lift ratio* maka lebih besar kekuatan asosiasinya (Santosa, 2007).

*Lift ratio* untuk aturan “X → Y” adalah perbandingan *confidence(X,Y)* dengan *benchmark confidence(X,Y)* yang didapatkan dari hasil perbandingan dari jumlah transaksi yang mengandung item Y terhadap jumlah transaksi pada basis data atau disebut juga *support(Y)*.

**2.2 Algoritma FP- Growth.**

*FP-Growth* menggunakan pendekatan yang berbeda dari paradigma yang selama ini sering digunakan, yaitu paradigma *Apriori*. Algoritma *Apriori* melakukan pemindaian basis data secara berulang – ulang untuk menentukan *frequent itemset*, sedangkan algoritma *FP-Growth* hanya memerlukan dua kali pemindaian basis data (Samuel, 2008).

Karakteristik algoritma *FP – Growth* adalah struktur data yang digunakan adalah

*tree*, yaitu disebut dengan *FP – tree*. Penggunaan *FP – tree* membuat algoritma *FP – Growth* dapat langsung mengekstrak *frequent itemset* (Erwin, 2009).

Algoritma *FP – Growth* menggunakan *divide* dan *conquer*, dimana membagi suatu permasalahan besar menjadi permasalahan – permasalahan yang lebih kecil (*divide*) dan pembagian dilakukan terus sampai ditemukan bagian masalah kecil yang mudah untuk dipecahkan (*conquer*) (Febriyana, 2009).

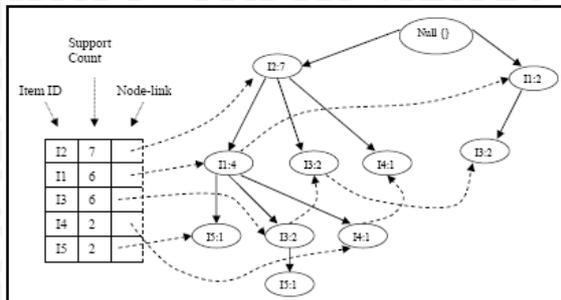
Fase dari algoritma *FP-growth* terdiri dari dua fase yaitu fase pembentukan *FP-tree* dan fase proses penggalian *FP-growth* (Yiapanis, 2006). *FP-tree* adalah struktur data berbentuk pohon (*tree*), yang terdiri dari : satu tabel *header*, satu *root* yang diberi label *null*, dan satu himpunan item *prefix subtree* sebagai node dari anak *root*. Masing-masing *entry* dalam tabel *header* adalah *frequent itemset*, dan setiap *record* terdiri dari dua atribut yaitu nama *item* (*item\_name*) dan *node* yang menjadi *head of node-link*. Sedangkan setiap node anak terdiri dari informasi nama *item*, *count*, dan *node-link*. Dimana *count* menunjukkan jumlah transaksi direpresentasikan oleh cabang yang mengandung node tersebut, dan *node-link* menghubungkan node ke node berikutnya pada *tree* tersebut yang mempunyai nama item yang sama atau *null* jika tidak ada (Maryeti,2006).

Algoritma *FP-growth* dimulai dengan dilakukan operasi pola *suffix* pada pola *frequent 1-itemset* dimana untuk masing – masing *suffix* dilakukan pembuatan *conditional pattern base*. *Conditional pattern base* merupakan suatu himpunan yang terdiri dari node yang membentuk lintasan awal (*prefix path*) yang berakhir pada *suffix* tertentu. Dari *conditional pattern base* tersebut akhirnya dibangun *conditional FP-tree* untuk melakukan penambahan secara rekursif. Pola yang dihasilkan didapatkan melalui kombinasi dari pola akhiran (*suffix*) dengan pola *frequent* yang dibangkitkan dari setiap *conditional FP-tree* dengan lintasan tunggal (Yiapanis,2006).

TID	Items
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3



T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3



### 2.3 Pengujian Akurasi

Pengujian akurasi *association rule* yang terbentuk yaitu dengan membandingkan *association rule* hasil proses data mining pada data latih dengan data hasil pada data uji. Penghitungan akurasi *association rule* dilakukan dengan menggunakan persamaan 2.7.

$$\text{Error} = \frac{|\text{support data uji} - \text{support data latih}|}{\text{support data latih}} \times 100\% \quad (2.7)$$

$$\text{akurasi} = 100\% - \text{error} \quad (2.8)$$

## III. ANALISA & PERANCANGAN SISTEM

### 3.1 Alur Proses Sistem

Data yang digunakan adalah data transaksi mobil pada dealer mobil dengan parameter jenis mobil, rentang umur, dan jenis kelamin. Nilai minimum *support* yang digunakan adalah 10% sedangkan nilai minimum *confidence* yang digunakan adalah 50%.

Alur proses dari Sistem ini adalah :

- 1) Tahap *preprocessing* dilakukan dengan menghapus kolom nama customer, tanggal transaksi dan tanggal lahir.
- 2) Tahap pembacaan pada data transaksi yang telah dilakukan *preprocessing* untuk menghitung frekuensi kemunculan tiap jenis mobil dan kombinasi jenis kelamin dengan rentang umur. Jenis kelamin dengan rentang umur dikombinasi atau dianggap satu kesatuan yang harus muncul bersamaan dalam kasus ini karena target aturan yang ingin didapatkan adalah asosiasi jenis kelamin dan rentang umur terhadap jenis mobil yang dibeli.
- 3) Tahap perhitungan *support* pada kandidat *frequent 1-itemset*, *support* kandidat *frequent 1-itemset* jenis mobil, *support*

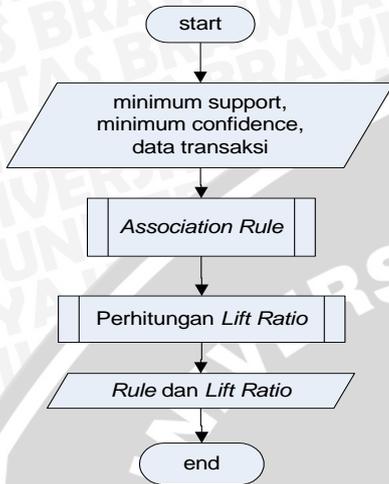
kandidat *frequent 1-itemset* kombinasi jenis kelamin dengan rentang umur.

- 4) Tahap pencarian *frequent 1-itemset* dari kandidat *frequent 1-itemset* baik jenis mobil maupun jenis kelamin dan rentang umur.
- 5) Tahap proses seleksi data transaksi yang memiliki baik *frequent 1-itemset* jenis mobil maupun *frequent 1-itemset* jenis kelamin dan rentang umur.
- 6) Tahap transformasi data dimana data hasil *preprocessing* data diubah bentuk sesuai dengan kebutuhan. Tahap transformasi yang dilakukan adalah mengubah bentuk tabel transaksi sehingga memudahkan proses *data mining*. Tabel diubah menjadi terdiri dua kolom yaitu id transaksi dan transaksi. Kolom id transaksi menyimpan informasi nomor id transaksi, sedangkan kolom transaksi menyimpan informasi transaksi yang berupa kombinasi jenis kelamin pembeli dengan rentang umur pembeli dan kode jenis mobil. Pengurutan item transaksi jenis mobil di belakang disebabkan karena target yang ingin diketahui adalah jenis mobil apa yang dibeli dengan karakteristik pembeli berdasarkan jenis kelamin pembeli dan rentang umur pembeli.
- 7) Tahap dimana setiap transaksi dari sampel data transaksi dibuat *FP-tree*. Pada pembuatan awal *root* diberikan nilai *null*. Pada *header table*, *head of table link* untuk semua item bernilai *null*.
- 8) Tahap pembacaan transaksi.
- 9) Tahap *conditional pattern base* yaitu mencari *suffix pattern* dan *prefix path* pada *FP-tree* yang telah terbentuk.
- 10) Tahap *conditional FP-tree* yaitu *support count* setiap itemset *prefix path* setiap *suffix* dijumlahkan. Item yang memiliki jumlah *support count* lebih besar atau sama dengan *minimum support* akan dibangkitkan dengan *conditional FP-tree*.
- 11) Tahap dimana didapatkan hasil *frequent itemset* dari setiap proses *conditional FP-tree* digabungkan menjadi satu dalam suatu tabel yaitu tabel *Frequent Itemset* kemudian dilakukan perhitungan nilai *confidence*.
- 12) *Frequent itemset* yang terpilih untuk dibangkitkan menjadi *rule*.
- 13) Selanjutnya *rule* yang terbentuk dilakukan uji kekuatan *rule* dengan menggunakan *lift ratio*. *Lift ratio*

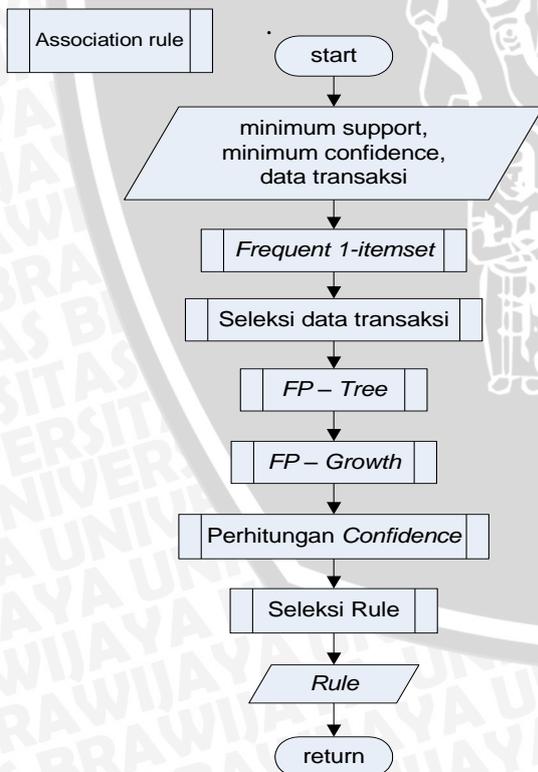
didapatkan dari hasil perbandingan antara *confidence* dengan *benchmark confidence*.

Analisis kebutuhan fungsional menjelaskan proses-proses yang terjadi di dalam aplikasi ini.

Proses-proses tersebut dijelaskan dalam *flowchart* diagram.

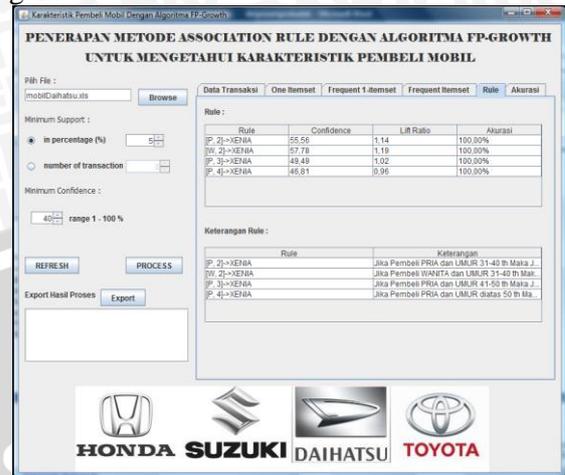


Gambar 3.1 *Flowchart* Perancangan Sistem



Gambar 3.2 *Flowchart* Proses Association Rule

Untuk tampilan program dapat dilihat dalam gambar berikut



Gambar 3.3 Tampilan Tabel *Rule* Setelah Uji Akurasi

#### IV. HASIL DAN PEMBAHASAN

##### 4.1 Pengujian Sistem

Pengujian pertama dilakukan untuk mengetahui tingkat kekuatan (*lift ratio*) *rule* yang dihasilkan. Hasil uji kekuatan (*lift ratio*) *rule* dapat dilihat pada table berikut:

Tabel 4.1 Detail Informasi *Rule* yang Bermanfaat dengan Algoritma *FP- Growth*.

Dealer	Rule	Support	Confidence	Lift Ratio
Daihatsu	[W, 2]->XENIA	5.37	53.66	1.11
	[P, 4]->XENIA	7.32	65.22	1.34
Toyota	[P, 2]->AVANZA 1300 CC	12.06	53.95	1.01
	[P, 3]->AVANZA 1300 CC	13.24	54.88	1.03
	[W, 3]->AVANZA 1300 CC	5.59	59.38	1.12
Suzuki	[P, 4]->PICK UP	6.3	56.25	1.21
Honda	[W, 1]->JAZZ GE8 E MT	5.97	80	1.45
	[P, 2]->JAZZ GE8 E MT	14.93	66.67	1.21
	[W, 2]->JAZZ GE8 E MT	13.43	64.29	1.16
	[P, 3]->JAZZ GE8 E MT	11.94	61.54	1.11
	[W, 3]->JAZZ GE8 E AT	7.46	55.56	2.19

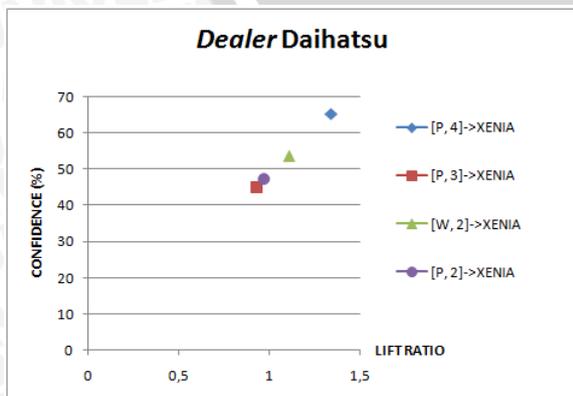
Pengujian kedua dilakukan untuk mengetahui tingkat akurasi dari *rule* yang dihasilkan masing-masing *dealer*. Hasil uji akurasi *rule* dapat dilihat pada table berikut:

Tabel 4.2 Akurasi *Rule* Hasil Uji Kekuatan *Lift Ratio*.

Dealer	Rule	Support Data uji	Support Data Latih	Akurasi
Daihatsu	[P,2]->XENIA	10.73	14.88	72.13%
	[P,3]->XENIA	9.76	9.76	100.00%
	[P,4]->XENIA	7.80	7.32	93.33%
	[W,2]->XENIA	4.88	5.37	90.91%
Toyota	[P, 2]->AVANZA 1300 CC	11.63	12.06	96.43%
	[P,3]->AVANZA 1300 CC	12.21	13.24	92.25%
	[W,3]->AVANZA 1300 CC	5.23	5.59	93.64%
	[P,4]->AVANZA 1300 CC	9.30	6.47	56.24%
Suzuki	[P, 2]->PICK UP	14.53	14.85	97.84%
	[P, 3]->PICK UP	10.06	10.92	92.05%
	[P, 4]->PICK UP	7.82	6.30	75.90%
Honda	[P, 2]->JAZZ GE8 E MT	11.43	14.93	76.57%
	[P, 3]->JAZZ GE8 E MT	14.29	11.94	80.36%
	[W, 1]->JAZZ GE8 E MT	2.86	5.97	47.86%
	[W, 2]->JAZZ GE8 E MT	14.29	13.43	93.65%
	[W, 3]->JAZZ GE8 E AT	8.57	7.46	85.14%

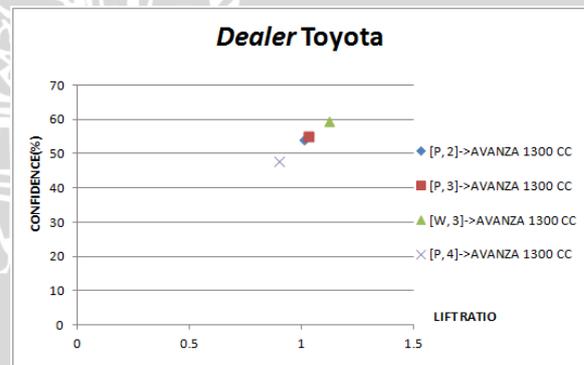
#### 4.2 Analisa Hasil

Berdasarkan hasil pengujian nilai kekuatan (*lift ratio*) *rule*, jumlah *rule* yang dihasilkan tiap *dealer* adalah sebagai *dealer* Suzuki 3 *rule*, *dealer* Honda 5 *rule*, *dealer* Toyota 4 *rule*, dan *dealer* Daihatsu 4 *rule*. Rata – rata nilai *lift ratio* yang didapatkan dari keseluruhan *rule* yang dihasilkan adalah 1.15. Nilai *lift ratio* tertinggi adalah 2.19 yang didapatkan dari *rule* hasil pengujian *dealer* Honda yaitu [W, 3]->JAZZ GE8 E AT atau jika pembeli wanita dan umur 41 sampai 50 tahun maka jenis mobil yang dibeli JAZZ GE8 E AT dengan tingkat kepercayaan sebesar 55.56%.



Gambar 4.1 Grafik *Confidence* dan *Lift Ratio* *Rule Dealer Daihatsu*

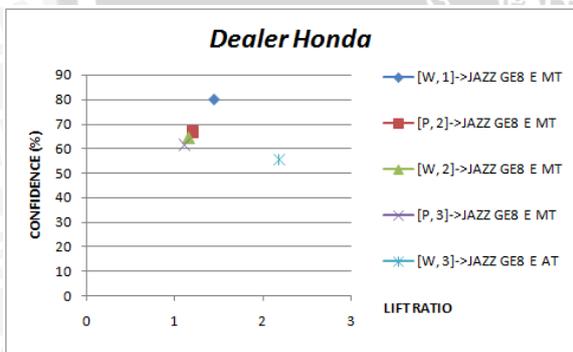
Pada Gambar 4.1 dapat dilihat grafik *confidence* dan *lift ratio* untuk *rule* yang dihasilkan dari pengujian *dealer* Daihatsu. Berdasarkan grafik tersebut, dari 4 macam *rule* yang dihasilkan *rule* yang memiliki nilai *lift ratio* lebih dari 1 adalah [P, 4]->XENIA dan [W, 2]->XENIA. Untuk *rule* [P, 4]->XENIA, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli pria yang berumur diatas 50 tahun sebesar 65.22% yang membeli jenis mobil Xenia dan tingkat kepercayaan pembeli pria yang berumur diatas 50 tahun masih lebih tinggi dari keseluruhan jumlah jenis mobil Xenia yang terjual. Untuk *rule* [W, 2]->XENIA, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli wanita yang berumur 31 sampai 40 tahun sebesar 53.66% yang membeli jenis mobil Xenia yang terjual dan tingkat kepercayaan pembeli pria yang berumur diatas 50 tahun masih lebih tinggi dari keseluruhan jumlah jenis mobil Xenia yang terjual. Maka dapat disimpulkan bahwa semakin tinggi nilai *confidence* suatu *rule* maka semakin tinggi pula nilai *lift ratio* *rule* tersebut, yang berarti bahwa semakin tinggi tingkat kepercayaan suatu *rule* maka nilai kekuatan *rule* tersebut juga semakin tinggi.



Gambar 4.2 Grafik *Confidence* dan *Lift Ratio* *Rule Dealer Toyota*

Pada Gambar 4.2 dapat dilihat grafik *confidence* dan *lift ratio* untuk *rule* yang dihasilkan dari pengujian *dealer* Toyota. Berdasarkan grafik tersebut, dari 3 macam *rule* yang dihasilkan *rule* yang memiliki nilai *lift ratio* lebih dari 1 [P, 2]->AVANZA 1300cc, [P, 3]->AVANZA 1300cc, dan [W, 3]->AVANZA 1300cc. Untuk *rule* [P, 2]->AVANZA 1300cc, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli pria

yang berumur 31 sampai 40 tahun sebesar 53.95% yang membeli jenis mobil Avanza 1300cc dan tingkat kepercayaan pembeli pria yang berumur 31 sampai 40 tahun masih lebih tinggi dari keseluruhan jumlah jenis mobil Avanza 1300cc yang terjual . Untuk *rule* [P, 3]->Avanza 1300cc, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli pria yang berumur 41 sampai 50 tahun sebesar 54.88% yang membeli jenis mobil Avanza 1300cc masih lebih tinggi dari keseluruhan jumlah jenis mobil Avanza 1300cc yang terjual. Untuk *rule* [W, 3]->Avanza 1300cc, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli wanita yang berumur 41 sampai 50 tahun sebesar 59.38% yang membeli Avanza 1300cc dan tingkat kepercayaan wanita yang berumur 41 sampai 50 tahun yang membeli jenis mobil Avanza 1300cc masih lebih tinggi dari keseluruhan jumlah jenis mobil Avanza 1300cc yang terjual. Maka dapat disimpulkan bahwa semakin tinggi nilai *confidence* suatu *rule* maka semakin tinggi pula nilai *lift ratio rule* tersebut, yang berarti bahwa semakin tinggi tingkat kepercayaan suatu *rule* maka nilai kekuatan *rule* tersebut juga semakin tinggi.

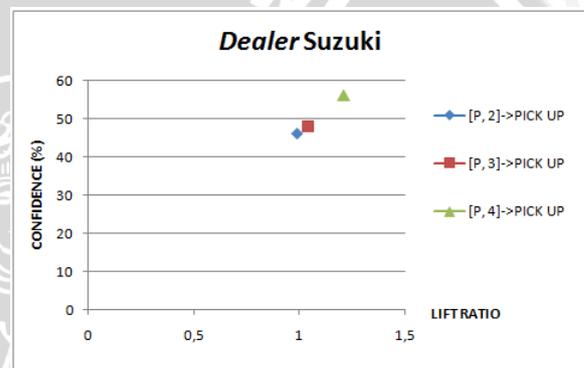


Gambar 4.3 Grafik Confidence dan Lift Ratio Rule Dealer Honda

Gambar 4.3 merupakan grafik *confidence* dan *lift ratio* untuk *rule* yang dihasilkan dari pengujian *dealer* Honda. Berdasarkan grafik tersebut, dari 5 macam *rule* yang dihasilkan *rule* yang memiliki nilai *lift ratio* lebih dari 1 adalah semua *rule*. Hal ini berarti tingkat kepercayaan pembeli wanita berumur 20 sampai 30 tahun dan 31 sampai 40 tahun yang membeli Jazz GE8 MT serta tingkat kepercayaan pembeli pria berumur 31 sampai 40 tahun dan 41 sampai 50 tahun yang membeli Jazz GE8 MT masih lebih tinggi

daripada keseluruhan jumlah Jazz GE8 MT yang terjual.

Selain itu, berdasarkan grafik Gambar 4.3 terdapat *rule* yang nilai *confidence*-nya tidak terlalu tinggi dibandingkan nilai *confidence rule* lain namun nilai *lift ratio* yang dimiliki justru sangat tinggi dibandingkan nilai *lift ratio rule* yang lain. Hal ini dijelaskan pada *rule* [W, 3]->JAZZ GE8 E AT dan menunjukkan bahwa frekuensi penjualan jenis mobil Jazz GE8 E AT yang tidak terlalu tinggi sehingga dengan tingkat kepercayaan yang tidak terlalu tinggi dapat menghasilkan kekuatan *rule* yang tinggi, namun tingkat kepercayaan *rule* tersebut masih tetap lebih tinggi daripada keseluruhan jumlah jenis mobil Jazz GE8 E AT yang terjual.



Gambar 4.4 Grafik Confidence dan Lift Ratio Rule Dealer Suzuki

Pada Gambar 4.4 dapat dilihat grafik *confidence* dan *lift ratio* untuk *rule* yang dihasilkan dari pengujian *dealer* Suzuki. Berdasarkan grafik tersebut, dari 3 macam *rule* yang dihasilkan *rule* yang memiliki nilai *lift ratio* lebih dari 1 adalah [P, 3]->PICK UP dan [P, 4]->PICK UP. Untuk *rule* [P, 3]->PICK UP, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli pria yang berumur 41 sampai 50 tahun sebesar 48.15% yang membeli jenis mobil Pick Up dan tingkat kepercayaan pembeli pria yang berumur berumur 41 sampai 50 tahun masih lebih tinggi dari keseluruhan jumlah jenis mobil Pick Up yang terjual. Untuk *rule* [P, 4]->PICK UP, hal ini menjelaskan bahwa dari keseluruhan jumlah pembeli wanita yang berumur diatas 50 tahun sebesar 56.25% yang membeli jenis mobil Pick Up yang terjual dan tingkat kepercayaan pembeli pria yang berumur diatas 50 tahun masih lebih tinggi dari keseluruhan jumlah jenis mobil Pick Up yang terjual.

Berdasarkan nilai *confidence* dan nilai *lift ratio rule* yang dihasilkan dalam pengujian ini, semakin tinggi nilai *confidence* suatu *rule* maka nilai *lift ratio rule* tersebut semakin tinggi pula dan semakin tinggi nilai frekuensi kemunculan item *consequent* suatu *rule* maka semakin kecil nilai *lift ratio rule* tersebut. *Rule* yang memiliki nilai *confidence* lebih tinggi daripada *rule* lain belum tentu memiliki nilai *lift ratio* yang lebih tinggi atau belum tentu memiliki nilai *lift ratio* lebih dari 1. Hal ini menjelaskan bahwa meskipun nilai kepercayaan (*confidence*) dari *rule* tersebut tinggi namun belum tentu *rule* tersebut bermanfaat karena nilai *confidence* menggambarkan tingkat kepercayaan item dalam *rule* tersebut muncul secara bersamaan terhadap jumlah keseluruhan item *antecedent* yang muncul sedangkan nilai *lift ratio* menggambarkan keseimbangan antara frekuensi kemunculan tiap item baik yang menjadi *antecedent* maupun *consequent* dalam *rule* tersebut secara independen dengan frekuensi kemunculannya secara bersamaan sehingga nilai *lift ratio* lebih dapat menunjukkan manfaat dari *rule* tersebut. Akan tetapi nilai minimum *support* yang digunakan juga mempengaruhi untuk mendapatkan *rule* yang benar-benar bermanfaat.

Pada pengujian akurasi, berdasarkan pada table 4.2 diperoleh akurasi *rule* yang terbentuk dari 4 *dealer*. Yang pertama *dealer* Daihatsu didapatkan bahwa *rule* [P,3]->XENIA memiliki akurasi paling tinggi, yaitu 100%. Hal ini menunjukkan bahwa pada data uji *dealer* Daihatsu sebanyak 204 data transaksi penjualan terdapat 20 data transaksi dengan jenis kelamin pria dan umur 41 sampai 50 tahun yang membeli jenis mobil Xenia dibandingkan dengan data latih sebanyak 409 data transaksi penjualan terdapat 40 data transaksi dengan jenis kelamin pria dan umur 41 sampai 50 tahun yang membeli jenis mobil Xenia. Sedangkan akurasi *rule* yang terbentuk dari *dealer* Toyota didapatkan bahwa *rule* [P,2]-> AVANZA 1300 CC memiliki akurasi paling tinggi, yaitu 96.43%. Hal ini menunjukkan bahwa pada data uji *dealer* Toyota sebanyak 170 data transaksi penjualan terdapat 20 data transaksi dengan jenis kelamin pria dan umur 31 sampai 40 tahun yang membeli jenis mobil AVANZA 1300 CC dibandingkan dengan data latih sebanyak 340 data transaksi penjualan terdapat 41 data transaksi dengan jenis kelamin pria dan umur

31 sampai 40 tahun yang membeli jenis mobil AVANZA 1300 CC. Untuk akurasi *rule* yang terbentuk dari *dealer* Suzuki didapatkan bahwa *rule* [P,2]->PICK UP memiliki akurasi paling tinggi, yaitu 97.84%. Hal ini menunjukkan bahwa pada data uji *dealer* Suzuki sebanyak 357 data transaksi penjualan terdapat 52 data transaksi dengan jenis kelamin pria dan umur 31 sampai 40 tahun yang membeli jenis mobil PICK UP dibandingkan dengan data latih sebanyak 714 data transaksi penjualan terdapat 106 data transaksi dengan jenis kelamin pria dan umur 31 sampai 40 tahun yang membeli jenis mobil PICK UP. Pada akurasi *rule* yang terbentuk dari *dealer* Honda didapatkan bahwa *rule* [W, 2]->JAZZ GE8 E MT memiliki akurasi paling tinggi, yaitu 93.65%. Hal ini menunjukkan bahwa pada data uji *dealer* Honda sebanyak 66 data transaksi penjualan terdapat 4 data transaksi dengan jenis kelamin wanita dan umur 31 sampai 40 tahun yang membeli jenis mobil [W, 2]->JAZZ GE8 E MT dibandingkan dengan data latih sebanyak 66 data transaksi penjualan terdapat 9 data transaksi dengan jenis kelamin pria dan umur 31 sampai 40 tahun yang membeli jenis mobil [W, 2]->JAZZ GE8 E MT.

Tingginya nilai akurasi *rule* menunjukkan bahwa tingginya penerapan *association rule* yang terbentuk pada data latih pada data uji. Semakin tinggi nilai akurasi *rule* maka semakin tinggi pula nilai kepercayaan bahwa *rule* tersebut menjadi pola asosiasi data karakteristik pembeli dengan jenis kelamin dan kelompok umur yang akan sering terjadi. Tingginya nilai akurasi *rule* memberikan informasi terkait kekuatan *rule* yang dihasilkan, yang bisa digunakan untuk bahan dasar prediksi pada pembelian mobil pada suatu *dealer* tertentu yang berdasarkan jenis kelamin dan umur pembeli.

## V. KESIMPULAN

Berdasarkan hasil analisis uji coba dapat diambil kesimpulan:

1. Dengan menerapkan algoritma *FP-Growth* pada data transaksi penjualan *dealer* mobil dengan karakteristik jenis kelamin dan jenis umur untuk jenis mobil tertentu di wilayah Mataram, didapatkan *rule* sebagai berikut :
  - a) Jenis Mobil Xenia *dealer* Daihatsu diminati oleh pembeli berjenis kelamin pria berumur diatas 50 tahun dan pembeli berjenis

- kelamin wanita berumur 31 sampai 40 tahun.
- b) Jenis Mobil Avanza *dealer* Toyota diminati oleh pembeli berjenis kelamin pria berumur 31 sampai 40 tahun dan pembeli berjenis kelamin wanita berumur 31 sampai 40 tahun.
  - c) Jenis Mobil Pick Up *dealer* Suzuki diminati oleh pembeli berjenis kelamin pria berumur diatas 50 tahun.
  - d) Jenis Mobil Jazz GE8 E MT *dealer* Honda diminati oleh pembeli berjenis kelamin pria berumur 31 sampai 50 tahun dan pembeli berjenis kelamin wanita berumur 20 sampai 40 tahun.
  - e) Jenis Mobil Jazz GE8 E AT *dealer* Honda diminati oleh pembeli berjenis kelamin wanita berumur 41 sampai 50 tahun.
2. *Rule* yang dihasilkan dari penerapan algoritma *FP-Growth* dalam permasalahan data transaksi penjualan *dealer* mobil di wilayah Mataram ini memiliki rata – rata tingkat kekuatan (*lift ratio*) sebesar 1.31 dengan tingkat kekuatan (*lift ratio*) *association rule* terbesar yang dihasilkan adalah 2.19 dan terkecil adalah 1.04 untuk *rule* dengan tingkat kepercayaan *rule* minimal 50%. Berdasarkan tingkat kekuatan (*lift ratio*) yang dihasilkan, dapat dikatakan bahwa *rule* yang bermanfaat tidak hanya dilihat berdasarkan nilai *confidence*, akan tetapi berdasarkan tiga parameter penting yaitu nilai *confidence*, nilai *lift ratio*, dan nilai minimum *support* yang digunakan.
  3. Akurasi *rule* yang dihasilkan dari penerapan *rule* hasil algoritma *FP-Growth* terhadap data uji, memiliki nilai rata – rata sebesar 83.28% dengan akurasi tertinggi sebesar 100% dan terkecil sebesar 47.86%. Nilai akurasi menunjukkan prosentase penerapan suatu *rule* pada data uji dari keseluruhan jumlah data transaksi pada data uji.

#### DAFTAR PUSTAKA

- Armstrong, Garry & Philip Kotler. 2004. Dasar-Dasar Pemasaran: Jilid 1 Edisi Kesembilan. Jakarta: PT.Indeks.
- Erwin. 2009. *Analisis Market Basket dengan Algoritma Apriori dan FP – Growth*. Jurnal Generic Volume 4 No. 2 Fakultas Ilmu Komputer.

Universitas Sriwijaya. Sumatera Selatan.

Febriyana, Fatma Rika. 2009. *Perbandingan Kecepatan dalam Pencarian Frequent Itemset antara Algoritma FP – Growth dan Cut Both Ways*. Universitas Brawijaya. Malang

[GAIKINDO] Gabungan Industri Kendaraan Bermotor Indonesia, Laporan Data periode 2005

Gilbert Mc Devinn, *Showroom Definition, Edisi 154, Penerbit Kompas Gramedia Building Unit 2 Lt 1, Jakarta, 2009.*

Gregorius S. Budhi, Leo W. Santoso, Karel R. Harisaputra. 2005. *Penggunaan Metode Data Mining CLARANS dan Naive Bayes Classification untuk Sistem Analisa Cluster dan Klasifikasi Data Terintegrasi*. Prosiding Sem Nas The Application of Technology Toward a Better Life 2005 buku 10.

Han, J dan Kamber, M. 2001. *Data Mining : Concepts and Techniques*. Morgan Kaufman. San Fransisco.

Han, J., Pie, J., Yin, Y. 2000. *Mining Frequent Patterns without Candidate Generation*, School of Computing Science Simon Fraser University.

Huda, Nuqson Masykur. 2010. *Aplikasi Data Mining untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa*. Semarang : Universitas Diponegoro.

Kantardzic, Mehmed. 2003 . *Data mining : Concepts, Models, Methods, and Algorithms*. John Wiley & Sons. New Jersey.

Kusrini, M.Kom. 2007. *Konsep dan Aplikasi Sistem Pendukung Keputusan*. CV Andi Offset. Yogyakarta.

Larose, Danniell T. *Discovering.2005. Knowledge in data : An Introduction to Data Mining*. John Wiley & Sons Inc. New Jersey.

Maryeti, Sri. 2006. *Analisis Perbandingan Algoritma FP-Growth dan algoritma Tree Projection dalam Pembangunan*

*Frequent Pattern.* STT Telkom Bandung.

Pramudiono, Iko. 2003. *Pengantar Data Mining : Menambang Permata Pengetahuan di Gudang Data*. <http://www.ilmukomputer.com>, tanggal akses 15 Maret 2011.

Rennolls, K. 2004. *Introduction to Data Mining – Chapter 2*.

Rochmah, Affriantari. 2010. *Perancangan Fitur Rekomendasi Film Website Solo Movie dengan Menggunakan Metode Algoritma Apriori*. Universitas Sebelas Maret. Surakarta.

Samuel, David. 2008. *Penerapan Struktur FP – Tree dan Algoritma FP – Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung.

Santosa, Budi. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Graha Ilmu. Yogyakarta.

Turban, E . 2005. *Decision Support System and Itelegence System*.Yogyakarta. Andi Ofset.

Verhein, Florian. 2008. *Frequent Pattern Growth (FP – Growth) Algorithm*. School of Information Technologies The University of Sydney. Australia.

Yiapanis, Paraskevas. 2006. *Parallel Mining of Minimal Sample Unique Itemsets*. School of Computer Science, University of Manchester.

Yulita, Marsela dan Veronica S. Moertini. 2002. *Analisis Keranjang Pasar dengan Algoritma Hash-Based pada Transaksi Penjualan di Apotek*. Integral Majalah Ilmiah Matematika dan Ilmu Pengetahuan Alam, Volume 9 No. 3. Jurusan Ilmu Komputer Universitas Katolik Parahyangan. Bandung.

