



PENGEMBANGAN ALGORITMA APRIORI UNTUK PENGAMBILAN KEPUTUSAN

THE DEVELOPMENT APRIORI ALGORITHM FOR DECISION- MAKING

¹Lismardiana, ²Herman Mawengkang, ³Erna Budhiarti Nababan
^{1,2,3}Magister S2 Fasilkom-TI Universitas Sumatera Utara Medan
¹kade_kadenta@yahoo.com

Diterima : 22 Juni 2015

Direvisi : 18 November 2015

Disetujui: 2 Desember 2015

ABSTRAK

Algoritma Apriori salah satu algoritma *data mining* dalam pembentukan *asosiasi rule mining*. Algoritma apriori adalah proses ekstraksi informasi dari suatu *database*, dilanjutkan dengan melakukan *frequent item/itemset* dan *candidate generation* dalam pembentukan *asosiasi rule mining* guna mendapatkan hasil nilai minimum *support* dan hasil nilai minimum *confidence*. Pada *database* yang cukup besar, algoritma apriori banyak menghasilkan *pattern frequent item/itemset* (pola sering muncul suatu item/itemset) yang banyak, karena harus melakukan *candidate generation* serta merekam *database* secara berulang-ulang. Dengan ini penulis berkeinginan mengembangkan algoritma apriori dengan melakukan penelitian tentang bagaimana meminimalkan *frequent item/itemset* pada apriori, tanpa melakukan *candidate generation* sehingga mempercepat tahapan penyelesaian pencarian *asosiasi rule mining*. Untuk solusi meminimalkan *frequent item/itemset* pada algoritma apriori, maka penulis menggunakan metode *FP-Growth*, dari hasil penelitian yang dilakukan dengan menggunakan dataset 1000 records pada *TransactionID-Sales*, pada apriori mulai dari k2, dihasil sebanyak 101 *frequent item/itemset*, sementara pada *FP-Growth* k2 sebanyak 40 *frequent item/itemset*. Dari jumlah hasil *frequent item/itemset* dapat disimpulkan bahwa dengan metode *FP-Growth* mampu meminimalkan jumlah *frequent item/itemset* pada algoritma apriori dan lebih efisien dari segi waktu, juga tahap penyelesaian lebih cepat, lebih terperinci dalam memaparkan hasil *frequent item/itemset* karena hasil *frequent* yang bernilai 1 masih diperhitungkan.

Kata Kunci : *Development Apriori, FP-Growth, Asosiasi Rule Mining, Frequent Item/itemset*

ABSTRACT

Apriori algorithm is one of the data mining algorithms in the formation of the association rule mining. Apriori algorithm is the process of extracting information from a database, followed by frequent item / itemset and candidate generation in the formation of the association rule mining in order to get the value of minimum support and minimum confidence value results. On the database is large enough, the algorithm generates a priori many frequent pattern item / itemset (pattern often appears in an item / itemsets) that much, because they have to

perform candidate generation and database to record repeatedly. By this author intends to develop a priori algorithm to conduct research on how to minimize frequent items / itemset on a priori, without candidate generation thereby accelerating the stage of completion of the search association rule mining. For solutions minimize frequent items / itemset in the algorithm a priori, the authors use the method of FP-Growth, the results of research conducted by using a dataset 1000 records on TransactionID-Sales, on a priori from k2, dihasil many as 101 frequent item / itemset, while at FP-Growth k2 40 frequent item / itemset. From the number of the results of frequent item / itemset can be concluded that the method of FP-Growth able to minimize the number of frequent item / itemset in the algorithm a priori and more efficient in terms of time, is also the stage of completion faster, more detailed in describing the results of frequent item / itemset as a result of frequent worth 1 is taken into account.

Keywords: Development Apriori, FP-Growth, Asosiasi Rule Mining, Frequent Item/itemset

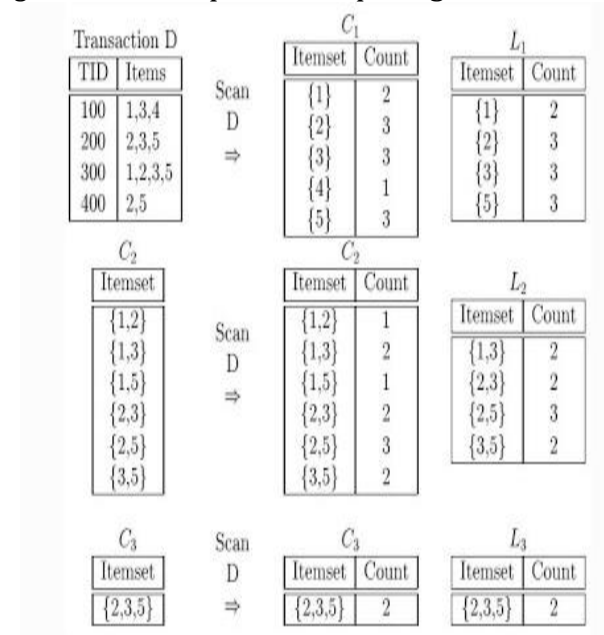
PENDAHULUAN

Data mining digunakan banyak tempat dan bidang penerapannya juga dapat bermacam macam, data mining mempelajari apa saja yang menjadi faktor utama dalam ketepatan sasaran pembelian suatu produk oleh konsumen. Kecerdasan bisnis merupakan proses pengubahan data menjadi informasi. Dari kumpulan informasi yang ada akan diambil polanya menjadi pengetahuan. Berry.M.J.A. dan Linoff.G.S.⁹ Tujuan kecerdasan bisnis ini adalah untuk mengubah data yang sangat banyak dan memiliki nilai bisnis melalui laporan analitik. Berry.M.J.A. dan Linoff.G.S.⁹

algoritma apriori salah satu algoritma data mining melakukan proses *ekstraksi informasi* pada *database* untuk menemukan aturan asosiasi antara suatu kombinasi *item/itemset*, Abdullah⁴. Penyelesaian masalah pada proses akstraksi informasi dari sebuah *database* atau data *mining* dengan melakukan proses generasi iterasi *frequent itemset* dalam jenis aturan asosiasi *rule mining (association rule mining)* sehingga menghasilkan nilai *support* dan *confidence*, Heena et al¹.

Pada *database* yang cukup besar proses pencairan asosiasi *rule mining* pada algoritma apriori membutuhkan waktu cukup lama, Moriwal² disebabkan semakin

besar database maka semakin banyak *frequent* yang dihasilkan dan *candidate generation*, dapat dilihat pada gambar 1.



Gambar 1. Proses iterasi asosiasi rule pada algoritma apriori

Sumber: Heena, 2014

Dari gambar 1 diatas diketahui bahwa semakin banyak / besar *database* akan semakin banyak timbul iterasi kombinasi *item/itemset* yang harus dilakukan setiap kali proses, sehingga mempengaruhi waktu penyelesaian pembentukan *asosiasi rule mining* dalam pencapaian nilai *support* dan *confidence*. Jadi inilah masalah algoritma

apriori sehingga perlu dikembangkan karena banyak menghasilkan *frequent item/itemset*.

Penelitian yang telah dilakukan berkaitan dengan algoritma apriori: Penelitian Kaur et al⁸ *Design and Implementation of Efficient Apriori Algorithm*, Heena et al¹ *Frequent Pattern Analysis of Moving Objects Using Apriori Algorithm International* dan Penelitian Kumar dan Rukmani³ *Implementation of Web Usage Mining Using Apriori and FP Growth Algorithms*.

Berdasarkan penelitian diatas, maka perlu dilakukan penelitian pada algoritma apriori untuk mengatasi munculnya *frequent item/itemset* dalam pencarian nilai *support* dan nilai *confidence* pada *database* yang cukup besar, sehingga dapat menghasilkan *asosiasi rule* mining tanpa melakukan *candidate generation*. Sehingga waktu penyelesaian pada proses *frequent item/itemset* pada algoritma apriori lebih efisien.

ALGORITMA APRIORI

Algoritma Apriori adalah salah satu algoritma pada data *mining* untuk mencari *frequent item/itemset* pada *transaksional database*. Algoritma apriori pertama kali diperkenalkan oleh R.Agarwal dan R Srikant untuk mencari *frequent* tertinggi dari suatu *database*, Kaur et al⁸. Penggunaan *bottom-up* pendekatan berulang. Untuk menentukan *asosiasi rule mining* sebuah *transaksi database*, diperlukan waktu dalam melakukan proses *frequent itemset*, menghasilkan kombinasi data yang cukup t banyak, Abdullah⁴. Proses ini dilakukan untuk mencari minimum nilai *support* dan minimum nilai *confidence*.

Algoritma apriori sangat mudah dipahami, tetapi ada beberapa kekurangan pada algoritma tersebut:

1. *Database Scanning*: *Database* transaksi perlu dipindai berulang kali untuk

menemukan *frequent itemset*. Jika ada n *item* dalam *database*, membutuhkan minimal n kali memindai *database*.

2. Pengaturan minimal *frequent item/itemset* untuk menentukan nilai *support* minimum.

3. Aturan *Asosiasi rule mining* dalam mendapatkan nilai minimum *confidence*

Langkah-langkah algoritma apriori sebagai berikut:

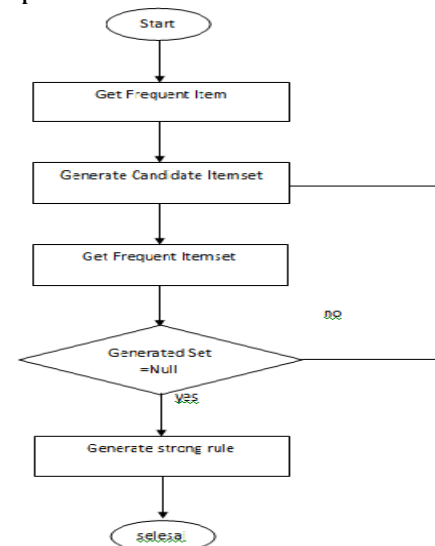
1. *Join*(penggabungan).

Pada proses ini setiap item dikombinasikan dengan item yang lainnya sampai tidak terbentuk kombinasi lagi.

2. *Prune*(pemangkasan).

Pada proses ini, hasil dari item yang telah dikombinasikan tadi lalu dipangkas dengan menggunakan minimum *support* yang telah ditentukan.

Dua proses utama tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset* pada algoritma Apriori.



Gambar 2. Deskripsi Algoritma Apriori

Pseudocode algoritma apriori:

Input:

D , a *database of transaction*; Min_Supp , the *minimum support count threshold*

Out put: L , *frequent itemsets in D*

Method:

```

L1 = find_frequent_1_itemsets(D);
For (k=2; Lk-1 ≠ ∅, k | i){
Ck=apriori_gen(L1-1);
For each transaction t ⊂ D{//scan D for counts
Ct = subset (Ck,t); // get the subsets for each candidate c ∈ Ct
        c.count / /;
}
Lk={c ⊂ Ck / c.counts ≥ min sup}
}
Return L = ∪k Lk;

```

ANALISIS ASOSIASI RULE MINING

Aturan asosiasi merupakan dalam data mining yang menemukan frequent itemset pada database. Asosiasi aturan data mining adalah mekanisme dalam data mining dalam aturan asosiasi, ekspresi implikasi dari bentuk $X \rightarrow Y$ di mana X adalah Y. Anteseden dan konsekuen ditetapkan item domain I. pendahuluan dan konsekuen adalah seperangkat item dari domain I. Dengan demikian $X \cap Y = \Phi$. Dukungan dari set item didefinisikan sebagai rasio jumlah transaksi yang mengandung item diatur pada jumlah total transaksi. Kepercayaan aturan asosiasi $X \rightarrow Y$ adalah probabilitas bahwa Y transaksi mengandung algoritma association rule mining X, Arora K. Rakesh dan Badal Dharmendra¹⁰ Rumus untuk mencari nilai support dan confidence adalah :

a. Support

$$\text{Support}(A \rightarrow B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Total Transaksi}}$$

b. Confidence

$$\text{Support}(A \rightarrow B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Total Transaksi}}$$

Analisis asosiasi didefinisikan suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk support (minimum support) dan

syarat minimum untuk confidence (minimum confidence).

FP-GROWTH

Mining tanpa melakukan candidate generation adalah teknik FP-Growth dengan menggunakan struktur data FP-tree, Han et al⁵. Dengan menggunakan cara ini scan database hanya dilakukan dua kali saja, tidak perlu berulang-ulang. Data akan direpresentasikan dalam bentuk FP-Tree. Setelah FP-Tree terbentuk, maka struktur data yang baik sekali untuk Frequent itemset akan diperoleh. Kumar B.S dan Rukmani .K.V.³ FP-Tree merupakan struktur data yang baik sekali untuk frequent Pattern mining, Han et al⁵ Struktur ini memberikan informasi yang lengkap untuk membentuk Frequent Pattern. Item-item yang tidak frequent (infrequent) sudah tidak ada dalam penggunaan FP-tree, Han et al⁵ Pembangunan FP-Tree dari sekumpulan data transaksi, akan diterapkan algoritma FP-Growth untuk mencari Frequent itemset yang signifikan, Han et al⁵. Algoritma FP-tree dibagi menjadi tiga langkah utama, yaitu:

1. Tahap Pembangkitan Conditional Pattern Base merupakan subdatabase yang berisi prefix path (lintasan e:1 prefix) dan pattern (pola akhiran). Pembangkitan conditionl pattern base didapatkan melalui FP-tree yang telah dibangun sebelumnya.
2. Tahap Pembangkitan Conditional FP-tree pada tahap ini, support count dari setiap item pada setiap conditional pattern base dijumlahkan, lalu setiap item yang memiliki jumlah support count lebih besar sama dengan minimum support count akan dibangkitkan dengan conditional FP-tree.
3. Tahap Pencarian frequent itemset apabila conditional FP-tree merupakan lintasan tunggal(single path), maka didapatkan frequent itemset dengan

melakukan kombinasi item untuk setiap *conditional FP-tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan *FP-growth* secara rekursif. Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapatkan *frequent itemset*.

Dengan menggunakan *FP-Growth*, kita dapat melakukan *Pattern Frequent itemset* dengan tidak membutuhkan waktu yang cukup lama.

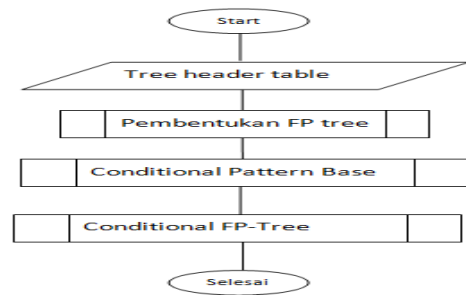
Pseudocode FP-Growth:

```

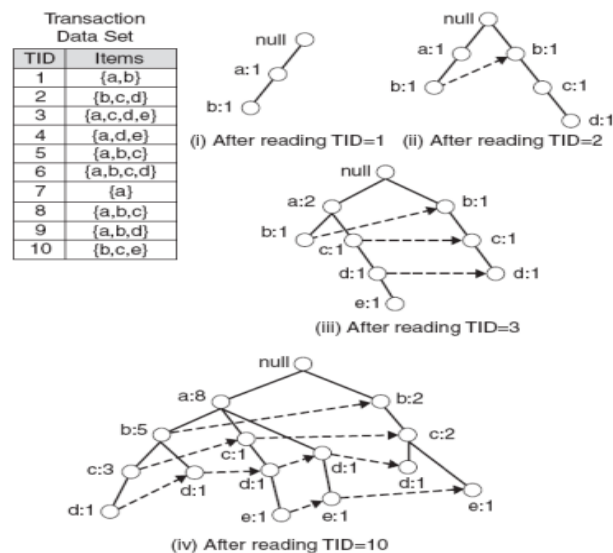
FP-Growth(Tree,  $\alpha$ )
for each( $a_i$  in the header of Tree) do {
 $\beta := a_i \cup \alpha$ 
generate( $\beta$  with support =  $a_i$ .support)
construct  $\beta$ 's conditional base pattern
and  $\beta$ 's conditional FP-Tree  $Tree_\beta$ 
if  $Tree_\beta \neq \emptyset$  then call FP-growth( $Tree_\beta, \beta$ )
    
```

Initially call:

FP-Growth(Tree, null)



Gambar 3. Deskripsi FP_Growth
Sumber :Han et al, 2000



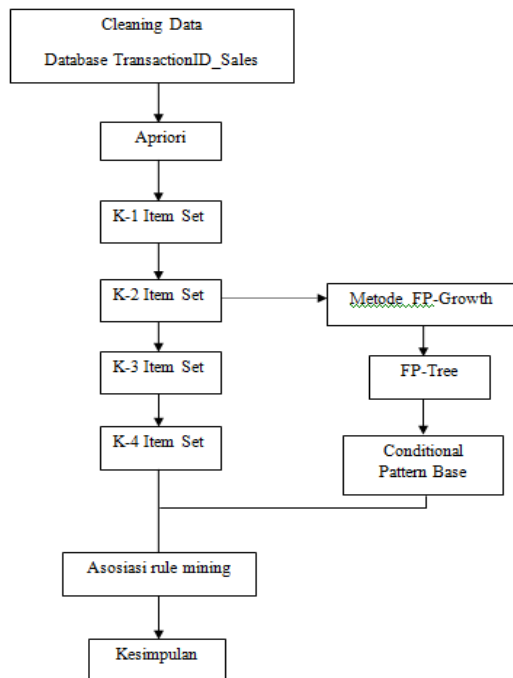
Gambar 4. Proses FP-Growth
Sumber: Han, 2000

METODE PENELITIAN

Teknik Pengembangan Algoritma Apriori

Pada penelitian ini dilakukan beberapa langkah sbb :

1. Penelitian ini dimulai dari pengumpulan data dan proses penyaringan data menggunakan Weka Explorer.
2. Tahap berikutnya dilakukan proses k-1 frequent item pada algoritma apriori
3. Hasil proses pada tahap k-1 frequent item pada apriori, untuk tahap ini digunakanlah



Gambar 5. Pengembangan Algoritma Apriori

Dari gambar 5 di atas pengembangan algoritma apriori dilaksanakan pada k1 item, pada apriori dilanjutkan dengan proses kombinasi ke 2 itemset (k-2 itemset) sedangkan dalam metode FP-Growth menggunakan FP-Tree.

Sumber dan Teknik Pengumpulan Data

Dalam melakukan penelitian ini, penulis menggunakan database yang bersumber dari: <https://support.spatialkey.com/spatialkey-sample-csv-data/> database yang berisikan transaksi penjualan sales di beberapa negara, database penelitian untuk apriori yang berformat .CSV.⁷ Database tersebut akan digunakan sebagai data pengujian untuk algoritma apriori dan metode FP-Growth dengan ketentuan sebagai berikut:

metode *FP-Growth* untuk meminimalkan *frequent item/itemset* dan mempersingkat cara kerja dalam pembentukan asosiasi *rule mining* untuk menghasilkan nilai *support* dan nilai *confidence*.

1. Data yang digunakan adalah *file database* berisikan *Transaction sales* penjualan *product* di beberapa negara dengan format. CSV
2. Banyak data yang digunakan adalah ± 1500 records.
3. *Atribute* yang digunakan dalam penelitian ini adalah : *City, State, Country*

Tabel 1. Data Spesifikasi untuk Penelitian

Nama Tabel	Data Transaksi Penjualan Sales
Jml Record	1000
Jml Transaksi	10.51
Jml Item	2997

Untuk *database* yang cukup besar, terjadi perekaman *database* secara berulang-ulang serta melakukan proses *candidate generation* secara bertahap-tahap.

HASIL PENELITIAN

Hasil Implementasi Algoritma Apriori untuk *Transaction ID_sales dataset 500 records*

Tabel 2. K-1 Apriori Frequent Item dataset 500 records

Frequent Item	Jlh.Support	Kombinasi	support
A2	2	A2	2
A3	2	A3	2
A4	1	B2	2
A10	1	B3	2
B2	2	B5	2
B3	2	C2	2
B4	1	C4	2
B5	2		
B6	1		
B7	1		
B9	1		
C2	2		
C4	2		
C8	1		
C21	1		

Dari tabel 2 di atas *frequent item* yang mempunyai nilai *support* 1 dieleminasi, dianggap tidak memenuhi nilai *support*. Pada generasi 1 hasil dari *frequent item* yang dihasilkan mempunyai nilai bobot 2, yang dijadikan sebagai *candidate generation* untuk *k-2 itemset*.

Tabel 3. K-2 Apriori Frequent Item dataset 500 records

Kombinasi	support	Kombinasi	support
A2,A3	2	A2,A3	2
A2,B2	2	A2,B2	2
A2,B3	2	A2,B3	2
A2,B5	2	A2,B5	2
A2,C2	2	A2,C2	2
A2,C4	2	A2,C4	2
A3,B2	2	A3,B2	2
A3,B3	2	A3,B3	2
A3,B5	2	A3,B5	2
A3,C2	2	A3,C2	2
A3,C4	2	A3,C4	2
B2,B3	2	B2,B3	2
B2,B5	2	B2,B5	2
B2,C2	2	B2,C2	2
B2,C4	2	B2,C4	2
B3,B5	2	B3,B5	2
B3,C2	2	B3,C2	2

Dari hasil kombinasi *k-2 frequent itemset* pada tabel 3 di atas menghasilkan 22 jumlah itemset.

Tabel 4 K-3 Apriori Frequent Itemset dataset 500 records

Kombinasi	Jumlah	Kombinasi	support
A2,A3,B2	2	A2,A3,B2	2
A2,A3,B3	2	A2,A3,B3	2
A2,A3,B5	2	A2,A3,B5	2
A2,A3,C2	2	A2,A3,C2	2
A2,A3,C4	2	A2,A3,C4	2
A3,B2,B3	2	A3,B2,B3	2
A3,B2,B5	2	A3,B2,B5	2
A3,B2,C2	2	A3,B2,C2	2
A3,B2,C4	2	A3,B2,C4	2
B2,B3,B5	2	B2,B3,B5	2
B2,B3,C2	2	B2,B3,C2	2
B2,B3,C4	2	B2,B3,C4	2
B3,B5,C2	2	B3,B5,C2	2
B3,B5,C4	2	B3,B5,C4	2
B5,C2,C4	2	B5,C2,C4	2

Dari hasil kombinasi *k-3 frequent itemset* pada tabel 4 di atas menghasilkan 15 jumlah itemset.

Tabel 5. K-4 Apriori Frequent Itemset dataset 500 records

Kombinasi	Jumlah	Kombinasi	support
A2,A3,B2,B3	2	A2,A3,B2,B3	2
A2,A3,B2,B5	2	A2,A3,B2,B5	2
A2,A3,B2,C2	2	A2,A3,B2,C2	2
A2,A3,B2,C4	2	A2,A3,B2,C4	2
A3,B2,B3,B5	2	A3,B2,B3,B5	2
A3,B2,B3,C2	2	A3,B2,B3,C2	2
A3,B2,B3,C4	2	A3,B2,B3,C4	2
B2,B3,B5,C2	2	B2,B3,B5,C2	2
B2,B3,B5,C4	2	B2,B3,B5,C4	2
B3,B5,C2,C4	2	B3,B5,C2,C4	2

Pada tabel 5 di atas dibuktikan bahwa *k-4 frequent itemset* sudah melampaui batas normal yang terjadi.

Tabel 6. Asosiasi Rule Mining untuk 500 records

itemset	Support	Min. Support 60%	Min. Confidence 70%
A2,A3,B2,B3	2	9%	9%
A2,A3,B2,B5	2	9%	9%
A2,A3,B2,C2	2	9%	9%
A2,A3,B2,C4	2	9%	9%
A3,B2,B3,B5	2	9%	9%
A3,B2,B3,C2	2	9%	9%
A3,B2,B3,C4	2	9%	9%
B2,B3,B5,C2	2	9%	9%
B2,B3,B5,C4	2	9%	9%
B3,B5,C2,C4	2	9%	9%

Pada tabel 6 di atas hasil asosiasi rule mining jika : Min. Support 60%, min.conf 70%, dengan hasil dari pencapaian rule mining masing-masing kelompok *itemset* menunjukkan tingkat prediksi 9%.

Hasil Implementasi Algoritma Apriori untuk Transaction ID_sales dataset 1000 records

Tabel 7. TransactionID_sales Dataset1000 records

TransactionID-Sales	Item
T250	A2,A3,A4,B2,B3,B5,B9,C2,C4,C8,C21
T500	A2,A3,A4,A6,B2,B3,B4,B5,B6,B7,B8,B9,B10,C2,C3,C4,C5,C6
T750	A2,A3,A4,A5,A6,B2,B3,B4,B5,B6,B7,B9,B14,B16,B23,C2,C3,C4,C5,C21
T1000	A2,A3,A10,B2,B3,B5,B4,B6,B7,C2,C3,C4

Tabel 8. K-1 Apriori Frequent Item dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A2	4	A2	4
A3	4	A3	4
A4	3	A4	3
B2	4	B2	4
B3	4	B3	4
B4	3	B4	3
B5	4	B5	4
B6	3	B6	3
B7	3	B7	3
B8	1	B9	3
B9	3	C2	4
B10	1	C3	3
B14	1	C4	4
B16	1	C5	2
B23	1	C21	2
C2	4		
C3	3		
C4	4		
C5	2		

dari tabel 8 diatas jumlah *frequent* pada *k-1* sebanyak 22 item.

Tabel 9. K-2 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A2,A3	4	A2,A3	4
A2,A4	3	A2,A4	3
A2,B2	4	A2,B2	4
A2,B3	4	A2,B3	4
A2,B4	3	A2,B4	3
A2,B5	4	A2,B5	4
A2,B6	3	A2,B6	3
A2,B7	3	A2,B7	3
A2,B9	3	A2,B9	3
A2,C2	4	A2,C2	4
A2,C3	3	A2,C3	3
A2,C4	4	A2,C4	4
A2,C5	2	A2,C5	2

dari tabel 9 k-2 apriori untuk dataset 1000 records menghasilkan 101 itemset. Hal ini membuktikan bahwa pada tahap ini dibutuhkan metode FP-Growth.

Tabel 10. K-3 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A2,A3,A4	3	A2,A3,A4	3
A2,A3,B2	4	A2,A3,B2	4
A2,A3,B3	4	A2,A3,B3	4
A2,A3,B4	3	A2,A3,B4	3
A2,A3,B5	4	A2,A3,B5	4
A2,A3,B6	3	A2,A3,B6	3
A2,A3,B7	3	A2,A3,B7	3
A2,A3,B9	3	A2,A3,B9	3
A2,A3,C2	4	A2,A3,C2	4
A2,A3,C3	2	A2,A3,C3	2
A2,A3,C4	4	A2,A3,C4	4
A2,A3,C5	2	A2,A3,C5	2
A3,A4,B2	3	A3,A4,B2	3
A3,A4,B3	3	A3,A4,B3	3
A3,A4,B4	3	A3,A4,B4	3
A3,A4,B5	3	A3,A4,B5	3
A3,A4,B6	2	A3,A4,B6	2
A3,A4,B7	2	A3,A4,B7	2
A3,A4,B9	3	A3,A4,B9	3
A3,A4,C2	3	A3,A4,C2	3
A3,A4,C3	2	A3,A4,C3	2
A3,A4,C4	3	A3,A4,C4	3
A3,A4,C5	2	A3,A4,C5	2
A4,B2,B3	3	A4,B2,B3	3
A4,B2,B4	2	A4,B2,B4	2
A4,B2,B5	3	A4,B2,B5	3
A4,B2,B6	2	A4,B2,B6	2
A4,B2,B7	2	A4,B2,B7	2
A4,B2,B9	3	A4,B2,B9	3

Dari tabel k-3 apriori menghasilkan jumlah 72 itemset.

Tabel 11. K-4 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A2,A3,A4,B2	3	A2,A3,A4,B6	2
A2,A3,A4,B3	3	A2,A3,A4,B7	2
A2,A3,A4,B4	3	A2,A3,A4,B9	2
A2,A3,A4,B5	3	A2,A3,A4,C3	2
A2,A3,A4,B6	2	A2,A3,A4,C5	2
A2,A3,A4,B7	2	A3,A4,B2,B4	2
A2,A3,A4,B9	2	A3,A4,B2,B6	2
A2,A3,A4,C2	3	A3,A4,B2,B7	2
A2,A3,A4,C3	2	A3,A4,B2,C3	2
A2,A3,A4,C4	3	A3,A4,B2,C5	2
A2,A3,A4,C5	2	A4,B2,B3,B4	2
A3,A4,B2,B3	3	A4,B2,B3,B6	2
A3,A4,B2,B4	2	A4,B2,B3,B7	2
A3,A4,B2,B5	3	A4,B2,B3,C3	2
A3,A4,B2,B6	2	A4,B2,B3,C5	2
A3,A4,B2,B7	2	B2,B3,B4,B5,	2
A3,A4,B2,B9	3	B2,B3,B4,B9	2
A3,A4,B2,C2	3	B2,B3,B4,C5	2
A3,A4,B2,C3	2	B3,B4,B5,B9	2
A3,A4,B2,C4	3	B3,B4,B5,C5	2
A3,A4,B2,C5	2	B4,B5,B6,B9	2
A4,B2,B3,B4	2	B4,B5,B6,C5	2
A4,B2,B3,B5	3	B5,B6,B7,B9	2
A4,B2,B3,B6	2	B5,B6,B7,C5	2

Tabel 12. K-4 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A4,B2,B3,B7	2	B6,B7,B9,C2	2
A4,B2,B3,B9	3	B6,B7,B9,C3	2
A4,B2,B3,C2	3	B6,B7,B9,C4	2
A4,B2,B3,C3	2	B6,B7,B9,C5	2
A4,B2,B3,C4	3	B7,B9,C2,C3	2
A4,B2,B3,C5	2	B7,B9,C2,C4	2
B2,B3,B4,B5	2	B7,B9,C2,C5	2
B2,B3,B4,B6	3	B9,C2,C3,C4	2
B2,B3,B4,B7	3	B9,C2,C3,C5	2
B2,B3,B4,B9	2	A2,A3,A4,B2	3
B2,B3,B4,C2	3	A2,A3,A4,B3	3
B2,B3,B4,C3	3	A2,A3,A4,B4	3
B2,B3,B4,C4	3	A2,A3,A4,B5	3
B2,B3,B4,C5	2	A2,A3,A4,C2	3
B3,B4,B5,B6	3	A2,A3,A4,C4	3
B3,B4,B5,B7	3	A3,A4,B2,B3	3
B3,B4,B5,B9	2	A3,A4,B2,B5	3
B3,B4,B5,C2	3	A3,A4,B2,B9	3
B3,B4,B5,C3	3	A3,A4,B2,C2	3
B3,B4,B5,C4	3	A3,A4,B2,C4	3
B3,B4,B5,C5	2	A4,B2,B3,B5	3
B4,B5,B6,B7	3	A4,B2,B3,B9	3
B4,B5,B6,B9	2	A4,B2,B3,C2	3
B4,B5,B6,C2	3	A4,B2,B3,C4	3
B4,B5,B6,C3	3	B2,B3,B4,B6	3
B4,B5,B6,C4	3	B2,B3,B4,B7	3
B4,B5,B6,C5	2	B2,B3,B4,C2	3
B5,B6,B7,B9	2	B2,B3,B4,C3	3
B5,B6,B7,C2	3	B2,B3,B4,C4	3
B5,B6,B7,C3	3	B3,B4,B5,B6	3
B5,B6,B7,C4	3	B3,B4,B5,B7	3
B5,B6,B7,C5	2	B3,B4,B5,C2	3
B6,B7,B9,C2	2	B3,B4,B5,C3	3
B6,B7,B9,C3	2	B3,B4,B5,C4	3
B6,B7,B9,C4	2	B4,B5,B6,B7	3
B6,B7,B9,C5	2	B4,B5,B6,C2	3
B7,B9,C2,C3	2	B4,B5,B6,C3	3
B7,B9,C2,C4	2	B4,B5,B6,C4	3
B7,B9,C2,C5	2	B5,B6,B7,C2	3
B9,C2,C3,C4	2	B5,B6,B7,C3	3
B9,C2,C3,C5	2	B5,B6,B7,C4	3

Pada tabel 12 di atas dari K-4 , (kombinasi item k-4) apriori menghasilkan jumlah 66 itemset.

Tabel 13. K-5 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A2,A3,A4,B2,B4	3	A2,A3,A4,B2,B3	4
A2,A3,A4,B2,B9	2	A2,A3,A4,B2,B5	4
A2,A3,A4,B2,C3	3	A2,A3,A4,B2,C2	4
A2,A3,A4,B2,B3	4	A2,A3,A4,B2,C4	4
A2,A3,A4,B2,B5	4	A3,A4,B2,B3,B5	4
A2,A3,A4,B2,B6	3	A3,A4,B2,B3,C2	4
A2,A3,A4,B2,C2	4	A3,A4,B2,B3,C4	4
A2,A3,A4,B2,C5	2	B2,B3,B4,B5,C4	4
A3,A4,B2,B3,B4	3	A2,A3,A4,B2,B4	3
A3,A4,B2,B3,B5	4	A2,A3,A4,B2,B6	3
A3,A4,B2,B3,B7	3	A2,A3,A4,B2,B7	3
A3,A4,B2,B3,B9	3	A2,A3,A4,B2,C3	3
A3,A4,B2,B3,C2	4	A3,A4,B2,B3,B4	3
A3,A4,B2,B3,C4	4	A3,A4,B2,B3,B6	3
A3,A4,B2,B3,C5	2	A3,A4,B2,B3,B7	3
A4,B2,B3,B4,B6	3	A3,A4,B2,B3,B9	3
A4,B2,B3,B4,B7	3	A3,A4,B2,B3,C3	3
A4,B2,B3,B4,B9	2	A4,B2,B3,B4,B5	3
A4,B2,B3,B4,C3	3	A4,B2,B3,B4,B6	3
A4,B2,B3,B4,C5	2	A4,B2,B3,B4,B7	3
B2,B3,B4,B5,B6	3	A4,B2,B3,B4,C2	3
B2,B3,B4,B5,B7	3	A4,B2,B3,B4,C3	3
B2,B3,B4,B5,C3	3	A4,B2,B3,B4,C4	3
B2,B3,B4,B5,C4	4	B2,B3,B4,B5,B6	3
B2,B3,B4,B5,C5	2	B2,B3,B4,B5,B7	3
A2,A3,A4,B2,B7	3	B2,B3,B4,B5,B9	3
A2,A3,A4,B2,C4	4	B2,B3,B4,B5,C2	3
A3,A4,B2,B3,B6	3	B2,B3,B4,B5,C3	3
A3,A4,B2,B3,C3	3	B3,B4,B5,B6,B7	3
A4,B2,B3,B4,B5	3	B3,B4,B5,B6,C2	3
A4,B2,B3,B4,C2	3	B3,B4,B5,B6,C3	3
A4,B2,B3,B4,C4	3	B3,B4,B5,B6,C4	3
B2,B3,B4,B5,B9	3	B4,B5,B6,B7,C2	3
B2,B3,B4,B5,C2	3	B4,B5,B6,B7,C3	3

Pada tabel 13 di atas dari K-5 , apriori menghasilkan 51 jumlah itemset.

Tabel 14. K-6 Apriori Frequent Itemset dataset 1000 records

Itemset	Jlh.Support	Kombinasi	support
A4,B2,B3,B4,B5,B6	3	A4,B2,B3,B4,B5,B6	3
A4,B2,B3,B4,B5,B7	3	A4,B2,B3,B4,B5,B7	3
A4,B2,B3,B4,B5,C2	3	A4,B2,B3,B4,B5,C2	3
A4,B2,B3,B4,B5,C3	3	A4,B2,B3,B4,B5,C3	3
A4,B2,B3,B4,B5,C4	3	A4,B2,B3,B4,B5,C4	3
B2,B3,B4,B5,B6,B7	3	B2,B3,B4,B5,B6,B7	3
B2,B3,B4,B5,B6,C2	3	B2,B3,B4,B5,B6,C2	3
B2,B3,B4,B5,B6,C3	3	B2,B3,B4,B5,B6,C3	3
B2,B3,B4,B5,B6,C4	3	B2,B3,B4,B5,B6,C4	3
B3,B4,B5,B6,B7,C2	3	B3,B4,B5,B6,B7,C2	3
B3,B4,B5,B6,B7,C3	3	B3,B4,B5,B6,B7,C3	3
B3,B4,B5,B6,B7,C4	3	B3,B4,B5,B6,B7,C4	3
A2,A3,A4,B2,B3,B9	2	A2,A3,A4,B2,B3,B9	2
A2,A3,A4,B2,B3,C5	2	A2,A3,A4,B2,B3,C5	2
A3,A4,B2,B3,B4,B9	2	A3,A4,B2,B3,B4,B9	2
A3,A4,B2,B3,B4,C5	2	A3,A4,B2,B3,B4,C5	2
A4,B2,B3,B4,B5,B9	2	A4,B2,B3,B4,B5,B9	2
A4,B2,B3,B4,B5,C5	2	A4,B2,B3,B4,B5,C5	2
B2,B3,B4,B5,B6,B9	2	B2,B3,B4,B5,B6,B9	2
B2,B3,B4,B5,B6,C5	2	B2,B3,B4,B5,B6,C5	2
B3,B4,B5,B6,B7,B9	2	B3,B4,B5,B6,B7,B9	2
B3,B4,B5,B6,B7,C5	2	B3,B4,B5,B6,B7,C5	2
B4,B5,B6,B7,B9,C2	2	B4,B5,B6,B7,B9,C2	2
B4,B5,B6,B7,B9,C3	2	B4,B5,B6,B7,B9,C3	2
B4,B5,B6,B7,B9,C4	2	B4,B5,B6,B7,B9,C4	2
B4,B5,B6,B7,B9,C5	2	B4,B5,B6,B7,B9,C5	2
B5,B6,B7,B9,C2,C3	2	B5,B6,B7,B9,C2,C3	2
B5,B6,B7,B9,C2,C4	2	B5,B6,B7,B9,C2,C4	2
B5,B6,B7,B9,C2,C5	2	B5,B6,B7,B9,C2,C5	2
B6,B7,B9,C2,C3,C4	2	B6,B7,B9,C2,C3,C4	2
B6,B7,B9,C2,C3,C5	2	B6,B7,B9,C2,C3,C5	2

Pada tabel 14 di atas dari K-6 , apriori menghasilkan 31 jumlah itemset.

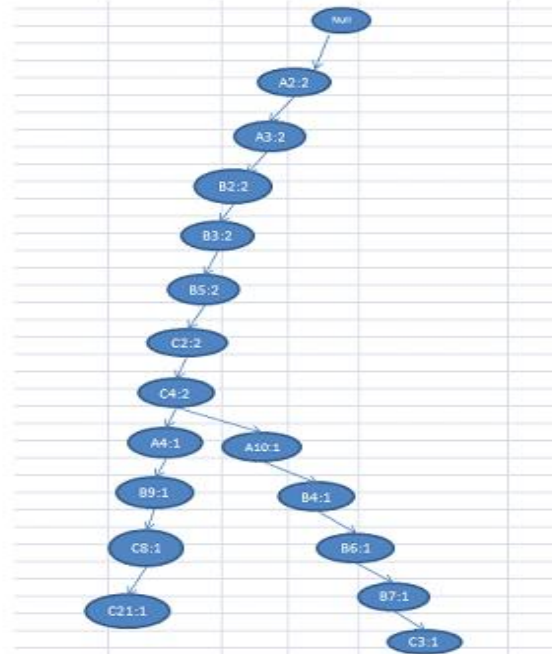
mining, sesuai dengan pendahuluan pada bab 1, bahwa apriori banyak menghasilkan candidate generation dan kombinasi frequent itemset.

Tabel 15. Asosiasi Rule Mining dataset 1000 records

Itemset	Support	Min. Support 60%	Min. Confidence 70%
A2,A3,A4,B2,B3,B5	4	5%	93%
A2,A3,A4,B2,B3,C2	4	5%	93%
A2,A3,A4,B2,B3,C4	4	5%	93%
A2,A3,A4,B2,B3,B4	3	4%	10%
A2,A3,A4,B2,B3,B6	3	4%	10%
A2,A3,A4,B2,B3,B7	3	4%	10%
A2,A3,A4,B2,B3,C3	3	4%	10%
A3,A4,B2,B3,B4,B5	3	4%	10%
A3,A4,B2,B3,B4,B6	3	4%	10%
A3,A4,B2,B3,B4,B7	3	4%	10%
A3,A4,B2,B3,B4,C2	3	4%	10%
A3,A4,B2,B3,B4,C3	3	4%	10%
A3,A4,B2,B3,B4,C4	3	4%	10%
A4,B2,B3,B4,B5,B6	3	4%	10%
A4,B2,B3,B4,B5,B7	3	4%	10%
A4,B2,B3,B4,B5,C2	3	4%	10%
A4,B2,B3,B4,B5,C3	3	4%	10%
A4,B2,B3,B4,B5,C4	3	4%	10%
B2,B3,B4,B5,B6,B7	3	4%	10%
B2,B3,B4,B5,B6,C2	3	4%	10%
B2,B3,B4,B5,B6,C3	3	4%	10%
B2,B3,B4,B5,B6,C4	3	4%	10%
B3,B4,B5,B6,B7,C2	3	4%	10%
B3,B4,B5,B6,B7,C3	3	4%	10%
B3,B4,B5,B6,B7,C4	3	4%	10%
A2,A3,A4,B2,B3,B9	2	3%	7%
A2,A3,A4,B2,B3,C5	2	3%	7%
A3,A4,B2,B3,B4,B9	2	3%	7%
A3,A4,B2,B3,B4,C5	2	3%	7%
A4,B2,B3,B4,B5,B9	2	3%	7%
A4,B2,B3,B4,B5,C5	2	3%	7%
B2,B3,B4,B5,B6,B9	2	3%	7%
B2,B3,B4,B5,B6,C5	2	3%	7%
B3,B4,B5,B6,B7,B9	2	3%	7%
B3,B4,B5,B6,B7,C5	2	3%	7%
B4,B5,B6,B7,B9,C2	2	3%	7%
B4,B5,B6,B7,B9,C3	2	3%	7%
B4,B5,B6,B7,B9,C4	2	3%	7%
B4,B5,B6,B7,B9,C5	2	3%	7%
B5,B6,B7,B9,C2,C3	2	3%	7%
B5,B6,B7,B9,C2,C4	2	3%	7%

Pada tabel 15 di atas adalah hasil dari asosiasi rule mining pada penelitian TransactionID_sales dataset 1000 records. Hasil asosiasi rule ada 3 bagian : bagian I: hasil frequent 4 dengan nilai support 5% dan nilai confidence 93% sebanyak 3 itemset lihat pada tabel 15 , bagian II : hasil frequent 3 dengan nilai support 4% dan nilai confidence 10% sebanyak 22 itemset lihat pada tabel 4.49, dan bagian III: hasil frequent 2 nilai support 3% dan nilai confidence 7% sebanyak 19 itemset.

Hasil Implementasi menggunakan Metode FP_Growth dengan dataset 500 records



Gambar 6. FP-tree dataset500

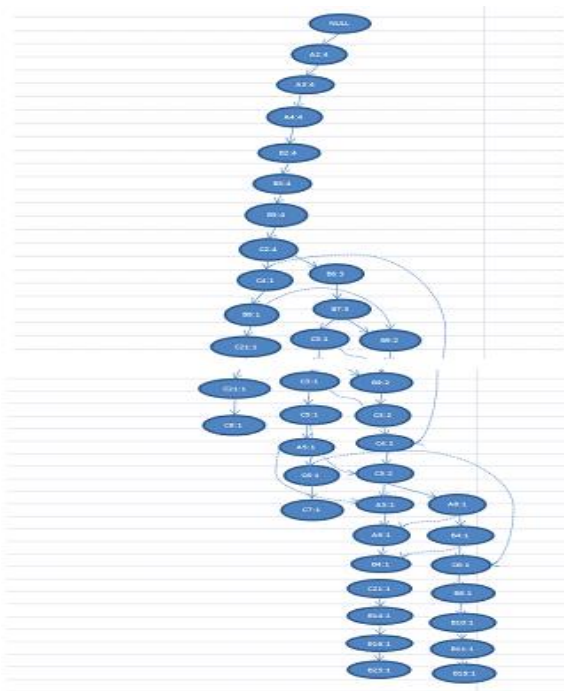
Pada gambar 6 di atas hasil frequent item dari FP-Growth menggunakan FP-tree. Jumlah frequent yang dihasilkan sebanyak 17 dengan 1 percabangan dengan dataset 500 records.

Tabel 16. Conditional Patternbase dataset 500

ItemID	Conditional pattern base	Conditional FP-Tree	Frequent pattern generated
T250	{{(A2,A3,B2,B3,B5,C2,C4):2}}	{{(A2,A3B2,B3,B5,C2,C4):2}<(A4,B9,C8,C21)}}	{{(A2,A3,B2,B3,B5,C2,C4):2}}
T500	{{(A2,A3,B2,B3,B5,C2,C4):2}}	{{(A2,A3B2,B3,B5,C2,C4):2}<(A4,B9,C8,C21)}}	{{(A2,A3,B2,B3,B5,C2,C4):2}}

Pada tabel 16 diatas masing-masing frequent pattern generation terdapat beberapa kelompok itemset : {(A2,A3,B2,B3,B5,C2,C4):2},{(A2,A3,B2,B3,B5,C2,C4):2}.

Hasil Implementasi menggunakan metode FP-Growth dengan dataset 1000 records



Gambar 7. FP-Tree dataset 1000

Pada gambar 7 di atas hasil frequent item dari FP-Growth menggunakan FP-tree. Jumlah frequent yang dihasilkan sebanyak 38 dengan 4 percabangan dengan dataset 1000 records.

Tabel 17. Tabel Conditional Pattern Base FP-Tree dataset 1000records

TID	Conditional pattern base	Conditional FP-Tree	Frequent pattern generated
T250	{(A2,A3,A4,B2,B3,B5,C2):4, (B4,B9,C21):2,(C8:1)}	{(A2,A3,A4,B2,B3,B5,C2):4, <(B4,B9,C21):2,<(C8:1)}	{(A2,A3,A4,B2,B3,B5,C2), (B4,B9,C21)}
T500	{(A2,A3,A4,B2,B3,B5,C2):4, (B6,B7,C3,C5):3,(C6:2),(C7: 1)}	{(A2,A3,A4,B2,B3,B5,C2):4,< (B6,B7,C3,C5):3,<(C6)2,<(C7: 1)}	{(A2,A3,A4,B2,B3,B5,C2), (B6,B7,C3,C5),(C6:2)}
T750	{(A2,A3,A4,B2,B3,B5,C2):4, (B6,B7,C3,C5):3,(C6:2),(C7: 1)}	{(A2,A3,A4,B2,B3,B5,C2):4, <(B6,B7,C3,C5):3,<(C6)2,<(C 7:1)}	{(A2,A3,A4,B2,B3,B5,C2), (B6,B7,C3,C5),(C6:2)}
T1000	{(A2,A3,A4,B2,B3,C2):4, (B6,B7,B9,C3,C4,C5):3,(A6, B4,C6):2,(B8,B10,B11,B18): 1}	{(A2,A3,A4,B2,B3,C2):4, <(B6,B7,B9,C3,C4,C5):3,<(A6, B4,C6):2,<(B8,B10,B11,B18): 1}	{(A2,A3,A4,B2,B3,C2), (B6,B7,B9,C3,C4,C5),(A6,B 4,C6)}

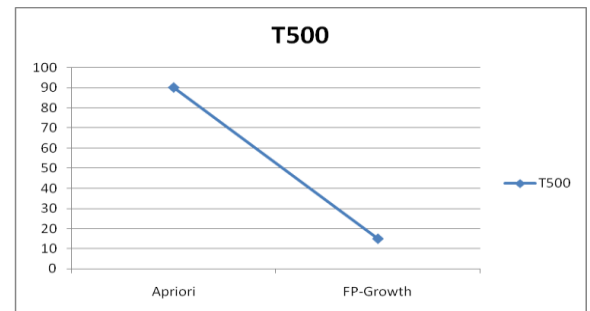
Hasil Analisis

Hasil penelitian penulis berkesimpulan bahwa jumlah *frequent item/itemset* yang dihasilkan dapat mempengaruhi waktu penyelesaian pembentukan *asosiasi rule mining*.

Waktu penyelesaian untuk dataset 500 records

Tabel 18. Perbandingan waktu Penyelesaian Apriori + FP Growth

TID	Apriori	FP-Growth
	Menit	Menit
T500	90	30

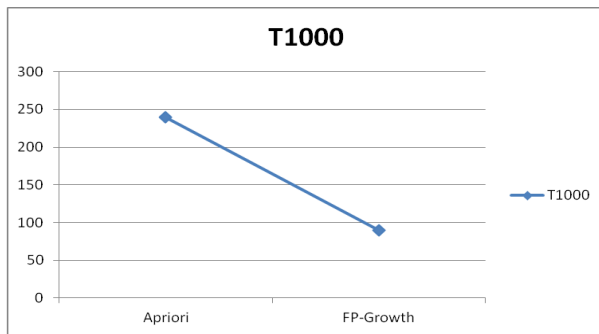


Gambar 8. Hasil Uji Coba Dataset 500 Records

Waktu penyelesaian untuk dataset 1000 records

Tabel 19. Penyelesaian Dataset 1000 Records

TID	Apriori	FP-Growth
	Menit	Menit
T500	240	90



Gambar 9. Hasil Uji Coba Dataset 1000 Records

Pada gambar 8 dan 9 perbandingan perbedaan penyelesaian apriori dengan *FP_Growth*. pada grafik diatas menunjukkan signifikan atas perbedaan waktu yang disebabkan dengan jumlah frequent yang dihasilkan, dimana jumlah hasil frequent yang sedikit lebih cepat menyelesaikan proses pembentukan asosiasi *rule mining*.

SIMPULAN

Berdasarkan percobaan yang telah dilakukan pada hasil penelitian bahwa: *FP-Growth* mampu meminimalkan jumlah *frequent item/itemset* pada algoritma *apriori*. pada *apriori* mulai dari *k2*, dihasilkan sebanyak 101 *frequent item/itemset*, sementara pada *FP-Growth k2* sebanyak 40 *frequent item/itemset*. Sehingga pemanfaatan dalam menggunakan *FP-Growth*: sedikit *frequent* yang dihasilkan maka waktu penyelesaian pembentukan *asosiasi rule mining* lebih cepat, lebih terperinci dalam memaparkan hasil *frequent item/itemset* karena hasil *frequent* yang bernilai 1 masih diperhitungkan.

DAFTAR PUSTAKA

¹Heena Rani, Shuchita Upadhyaya dan Vinod Kumar.2014. *Frequent Pattern Analysis of Moving Objects Using Apriori Algorithm* International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 Volume-3, Issue-4 April .

²Moriwal. Rahul .2014. *FP-growth Tree for large and Dynamic Data Set and Improve Efficiency*. ISSN 1746-7659, England, UK Journal of Information and Computing Science Vol. 2:083-090

³Kumar B.S dan Rukmani .K.V. 2010. *Implementation of Web Usage Mining Using APRIORI and FP Growth Algorithms*. Int. J. of Advanced Networking and Applications 400 Volume:01, Issue:06, Pages: 400-404

⁴Abdullah Saad Almalaise Alghamdi. *Efficient Implementation of FP Growth Algorithm-Data Mining on Medical Data*. IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.12, December 2011

⁵Han et al 2000. *Mining Frequent Patterns without Candidate Generation* In Proceedings of the 2000 ACM SIGMOD international Conference on Management of Data (Dallas, Texas, United States, May 15 -18, 2000). SIGMOD '00. ACM Press, New York, NY, 1-12.

⁶Tanna P & Ghodasara Y.2014. *Using Apriori with WEKA for Frequent Pattern Mining* International Journal of Engineering Trends and Technology (IJETT) - Volume 12 Number 3 - Jun 2014

⁷(<https://support.spatialkey.com/spatialkey-sample-csv-data/> data sales di download pada tanggal 18 Juni 2015 pukul 22.45 WIB)

⁸Kaur et al 2014. *Design and Implementation of Efficient Apriori Algorithm* International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 5 ISSN: 2321-8169 1205- 1208

⁹Berry.M.J.A. & Linoff.G.S. 2004. *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management* Second Edition .

¹⁰Arora K. Rakesh dan Badal Dharmendra. 2014. *Mining Association Rules to Improve Academic Performance* ", International Journal of Computer Science and Mobile Computing, Vol. 3 : 1