

SKRIPSI

METODE ASSOCIATION RULE DALAM MENGANALISA POLA BELANJA KONSUMEN PADA DATA TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA FP-GROWTH

METHOD OF ASSOCIATION RULE IN ANALYZING CONSUMER SPENDING PATTERN ON SALE TRANSACTION DATA USING FP-GROWTH ALGORITM

Diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Teknik
Informatika



Disusun Oleh:

NAMA : Dessy Chaerunnissa
NIM : A11.2011.06130
Program Studi : Teknik Informatika – S1

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO
SEMARANG**

2015

PERSETUJUAN SKRIPSI

Nama : Dessy Chaerunnissa
Nim : A11.2011.06130
Program Studi : Teknik Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Metode Association Rule Dalam Menganalisa Pola
Belanja Konsumen Pada Data Transaksi Penjualan
Menggunakan Algoritma FP-Growth

Tugas Akhir ini telah diperiksa dan disetujui,
Semarang, 9 Juli 2015

Menyetujui,

Dosen Pembimbing

Dekan Fakultas Ilmu Komputer

Edy Mulyanto, S.Si, M.Kom
NIP. 0686.11.1993.040

Dr. Drs Abdul Syukur, MM
NIP. 0686.11.1992.017

PENGESAHAN DEWAN PENGUJI

Nama : Dessy Chaerunnissa
Nim : A11.2011.06130
Program Studi : Teknik Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Metode Association Rule Dalam Menganalisa Pola Belanja Konsumen Pada Data Transaksi Penjualan Menggunakan Algoritma FP-Growth

Tugas Akhir ini telah disetujui dan dipertahankan dihadapan Dewan Penguji pada Sidang tugas akhir tanggal 9 Juli 2015. Menurut pandangan kami, tugas akhir ini memadai dari segi kualitas maupun kuantitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom.)

Semarang, 9 Juli 2015

Dewan Penguji :

Catur Supriyanto, S.Kom, M. CS

Anggota 1

Setia Astuti, S.Si, M.Kom

Anggota 2

Hanny Haryanto, S.Kom, M.T

Ketua Penguji

PERNYATAAN KEASLIAN SKRIPSI

Sebagai mahasiswa Universitas Dian Nuswantoro, yang bertanda tangan dibawah ini, saya:

Nama : Dessy Chaerunnissa

NIM : A11.2011.06130

Menyatakan bahwa karya ilmiah saya yang berjudul :

Metode Association Rule Dalam Menganalisa Pola Belanja Konsumen Pada Data Transaksi Penjualan Menggunakan Algoritma FP-Growth

Merupakan karya asli saya (kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya). Apabila di kemudian hari, karya saya disinyalir bukan merupakan karya asli saya, yang disertai bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar saya beserta hak dan kewajiban yang melekat pada gelar tersebut. Demikian surat pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Semarang

Pada tanggal : 9 Juli 2015

Yang menyatakan

(Dessy Chaerunnissa)

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai mahasiswa Universitas Dian Nuswantoro, yang bertanda tangan di bawah ini, saya :

Nama : Dessy Cherunnissa

Nim : A11.2011.06130

Demi mengembangkan Ilmu Pengetahuan, menyetujui untuk memberikan kepada Universitas Dian Nuswantoro Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul :

Metode Association Rule Dalam Menganalisa Pola Belanja Konsumen Pada Data Transaksi Penjualan Menggunakan Algoritma FP-Growth .

Beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Dian Nuswantoro berhak menyimpan, mengcopy ulang (memperbanyak), menggunakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan menampilkan/mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Dian Nuswantoro, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian surat pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Semarang

Pada tanggal: 9 Juli 2015

Yang menyatakan

(Dessy Chaeunnissa)

UCAPAN TERIMA KASIH

Dengan memanjatkan puji syukur kehadiran Allah SWT, yang telah melimpahkan rahmat karunia-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “**Metode Association Rule Dalam Menganalisa Pola Belanja Konsumen Pada Data Transaksi Penjualan Menggunakan Algoritma FP-Growth.**” pada waktu yang telah ditetapkan karena dukungan dari pihak yang tidak ternilai besarnya. Oleh karena itu penulis mengucapkan rasa terima kasih kepada :

1. Dr. Ir. Edi Noersasongko, M.kom, selaku rektor Universitas Dian Nuswantoro.
 2. Dr. Drs. Abdul Syukur, MM, selaku Dekan Fakultas Ilmu Komputer.
 3. Dr. Heru Agus Santoso, M.Kom, selaku Ketua Program Studi Teknik Informatika – S1.
 4. Edy Mulyanto, S.Si, M.Kom selaku pembimbing tugas akhir yang telah memberikan bimbingan dan masukan kepada penulis dalam menyusun laporan tugas akhir hingga selesai.
 5. Ardytha Luthfiarta, M.Kom, M.Cs, dan Junta Zeniarna M.Kom, M.CS atas petunjuk, saran, penjelasan, dan pengarahan yang diberikan sehubungan dengan penelitian yang penulis lakukan.
 6. Dosen-dosen pengampu di Fakultas Ilmu Komputer Teknik Informatika Universitas Dian Nuswantoro Semarang yang telah memberikan ilmu dan pengalamannya sehingga penulis dapat menerapkan ilmu yang telah disampaikan.
 7. Orang tua tercinta yang telah memberikan dukungan baik spiritual maupun material, sehingga penulis dapat menyelesaikan penelitian.
 8. Teman-teman dan seluruh kerabat yang telah memberikan dukungan dalam bentuk apapun sehingga dapat membantu penulis menyelesaikan penelitian
- Semoga Tuhan Yang Maha Esa memberikan balasan yang lebih besar kepada beliau-beliau. Penulis berharap penelitian ini dapat lebih disempurnakan sehingga lebih berkembang dan dapat bermanfaat sebagaimana fungsinya

Semarang, Juli 2015

Penulis

ABSTRAK

Banyaknya data transaksi yang tersimpan dalam database menyebabkan penumpukan data. Data tersebut dapat dimanfaatkan untuk diolah lebih lanjut menjadi suatu informasi yang dapat digunakan sebagai acuan bagi pihak Swalayan untuk menentukan pengambilan keputusan dalam kebijakan dan strategi bisnis. Dengan adanya data mining diharapkan dapat membantu Swalayan Gelael Candi Semarang untuk menggali informasi yang terkandung didalam data transaksi menjadi sebuah pengetahuan (knowledge) yang baru. Metode yang digunakan yaitu Market Basket Analysis. Association Rule, yaitu prosedur dalam Market Basket Analysis untuk mencari pengetahuan berupa hubungan antar item dalam satu dataset dan menampilkanya dalam bentuk pola asosiasi yang menjelaskan tentang kebiasaan konsumen dalam berbelanja. Algoritma yang digunakan untuk membantu menemukan pola asosiasi yaitu algoritma Frequent Pattern Growth (FP-Growth), merupakan algoritma yang sangat efisien dalam pencarian frequent itemset dalam sebuah kumpulan data dengan membangkit strukur prefix-tree atau disebut dengan FP-Tree. Pola asosiasi ditentukan oleh dua parameter, yaitu support (nilai penunjang) dan confidence (nilai kepastian). Salah satu pola yang dihasilkan dari analisis terhadap data transaksi bulan Oktober 2014 yaitu jika membeli salties maka membeli soft drink dengan nilai support = 10.27% dan nilai confidence = 45,6% yang merupakan pola dengan nilai support dan confidence tertinggi.

Kata kunci : *Data Mining, Market Basket Analysis, Association Rule, FP-Growth*

DAFTAR ISI

HALAMAN JUDUL.....	i
PERSETUJUAN SKRIPSI	ii
PENGESAHAN DEWAN PENGUJI.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	v
UCAPAN TERIMA KASIH.....	vi
ABSTRAK	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II LANDASAN TEORI	6
2.1 Tinjauan Studi	6
2.2 Landasan Teori	10
2.2.1 Definisi Data Mining.....	10
2.2.2 Tahap-Tahap Data Mining	12
2.2.3 Teknik Data Mining	14
2.2.4 <i>Market Basket Analysis</i>	17
2.2.5 <i>Association Rule</i>	18
2.2.6 Algoritma FP-Growth	21
2.2.6.1. Pembuatan <i>FP-Tree</i>	22
2.2.6.2. Penerapan Algoritma FP-Growth	24
2.2.7 CRISP-DM.....	24

2.3	Kerangka Pemikiran	27
BAB III METODE PENELITIAN.....		28
3.1	Instrumen Penelitian	28
3.2	Objek Penelitian	28
3.3	Sumber Data	28
3.4	Metode Pengumpulan Data	29
3.5	Tahap Penelitian	29
3.5.1.	Pengumpulan Data	29
3.5.2.	Penyeleksian Data	30
3.5.3.	Preprocessing/Cleaning.....	30
3.5.4.	Transformasi Data.....	35
3.5.5.	Data Mining	37
3.5.6.	Interpretation/Evaluation.....	37
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Pengumpulan Data	38
4.2	Pengumpulan Data	38
4.3	Preprocessing/Cleaning	40
4.4	Transformasi Data	40
4.5	Data Mining.....	40
4.6	Implementasi Java	62
4.7	Interpretation dan Evaluation	67
BAB V KESIMPULAN DAN SARAN.....		68
5.1.	Kesimpulan.....	68
5.2.	Saran	68
DAFTAR PUSTAKA		69

DAFTAR TABEL

Tabel 2.1 <i>State of the art</i> penelitian dengan metode <i>association rules</i>	8
Tabel 3.1 Penjelasan atribut dalam dataset	30
Tabel 3.2 Sample data yang akan diolah.....	31
Tabel 3.3 Data transaksi dengan nama barang	33
Tabel 3.4 Dataset hasil transformasi	36
Tabel 4.1 Contoh 20 data transaksi	41
Tabel 4.2 Frequent itemset	42
Tabel 4.3 Frequent list dengan min <i>support</i> $\geq 0,15$	43
Tabel 4.4 Transaksi yang sudah diurutkan berdasarkan F-List.....	44
Tabel 4.5 Hasil Frequent Itemset	60
Tabel 4.6 Tabel subset yang layak	61

DAFTAR GAMBAR

Gambar 2.1 Tahap proses data mining.....	12
Gambar 2.2 Metodologi <i>Association Rule</i>	21
Gambar 2.3 Diagram alir algoritma <i>FP-Growth</i>	23
Gambar 2.4 Proses Data Mining Menurut CRISP-DM	25
Gambar 3.1 <i>Sample dataset</i> transaksi penjualan.....	29
Gambar 4.1 Hasil pembentukan TID 1	45
Gambar 4.2 Hasil pembentukan TID 2	45
Gambar 4.3 Hasil pembentukan TID 3	46
Gambar 4.4 Hasil pembentukan TID 4	47
Gambar 4.5 Hasil pembentukan TID 5	47
Gambar 4.6 Hasil pembentukan TID 6	48
Gambar 4.7 Hasil pembentukan TID 7	49
Gambar 4.8 Hasil pembentukan TID 10	50
Gambar 4.9 Hasil pembentukan TID 11	50
Gambar 4.10 Hasil pembentukan TID 12	51
Gambar 4.11 Hasil pembentukan TID 14	52
Gambar 4.12 Hasil pembentukan TID 15	52
Gambar 4.13 Hasil pembentukan TID 17	53
Gambar 4.14 Hasil pembentukan TID 18	54
Gambar 4.15 Hasil pembentukan TID 19	54
Gambar 4.16 Hasil pembentukan FP-Tree item 87.....	55
Gambar 4.17 Hasil pembentukan FP-Tree item 75.....	56
Gambar 4.18 Hasil pembentukan FP-Tree item 20.....	57
Gambar 4.19 Hasil pembentukan FP-Tree item 76.....	57
Gambar 4.20 Hasil pembentukan FP-Tree item 5.....	58
Gambar 4.21 Hasil pembentukan FP-Tree item 81.....	59
Gambar 4.22 Hasil pembentukan FP-Tree item 21.....	59
Gambar 4.23 Hasil pembentukan FP-Tree item 7.....	60
Gambar 4.24 Tampilan menu utama.....	63

Gambar 4.25 Tampilan input file	63
Gambar 4.26 Tampilan output file	64
Gambar 4.27 Tampilan input min support dan min confidence.....	65
Gambar 4.28 Tampilan saat running program	65
Gambar 4.29 Tampilan hasil pada notepad.....	66
Gambar 4.30 Hasil aturan asosiasi yang terbentuk	66

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pasar swalayan merupakan pasar modern yang menyediakan berbagai kebutuhan konsumen, misalnya kebutuhan primer dan sekunder. Fenomena pasar modern ini telah mengalami perkembangan pesat dan juga menjamur di kota-kota besar Indonesia hingga ke pelosok desa, bahkan seiring dengan maraknya pasar-pasar modern membuat pasar tradisional mulai tergeserkan. Pola pikir konsumenpun mulai beralih dari pasar tradisional dengan orientasi harga murah kini berpaling pada pasar-pasar modern yang lebih memperhatikan aspek keamanan, kenyamanan, kebersihan, keramahan pelayanan, kelengkapan barang, dan penempatan rak barang. Semarang merupakan salah satu kota dimana terdapat banyak bangunan-bangunan swalayan, salah satunya adalah Swalayan Gelael Candi.

Swalayan selalu berusaha mewujudkan kepuasan bagi konsumen dengan menyediakan produk yang berkualitas, layanan yang unggul, dan akrab bersahabat, serta dalam suasana belanja yang menyenangkan, namun karena terjadi persaingan dengan swalayan lainnya maka diperlukan strategi-strategi untuk mempertahankan bisnis tersebut [1]. Pasar swalayan pasti mengalami naik turunnya peningkatan penjualan barang.

Salah satu strategi yang dapat digunakan untuk meningkatkan volume penjualan adalah dengan memahami dan mengetahui pola belanja konsumen [2]. Pengetahuan tentang pola belanja tersebut kemudian dapat dimanfaatkan untuk pengambilan keputusan dalam strategi bisnis guna meningkatkan pendapatan swalayan. Data yang digunakan untuk menghasilkan suatu informasi mengenai pola belanja konsumen yaitu data transaksi penjualan

yang telah terjadi. Analisa ini menghasilkan informasi mengenai keterkaitan antar barang yang sering dibeli secara bersamaan dalam satu transaksi.

Salah satu cara yang efektif untuk mengoptimalkan pencarian pola keterkaitan antar barang yang dibeli secara bersamaan dalam satu transaksi yang dilakukan pelanggan dalam pembelian beberapa produk yaitu metode asosiasi atau *association rules* [3]. Metode asosiasi merupakan suatu teknik dalam *Data Mining* untuk menentukan aturan asosiasi hubungan antar item barang. Pengetahuan mengenai asosiasi antar jenis item, jenis-jenis item yang muncul bersamaan pada tiap transaksi, dapat menjadi masukan penting dalam melakukan usaha peningkatan penjualan. Misalkan dalam melakukan promosi untuk pembelian satu jenis item dapat membeli jenis item lain dengan harga lebih murah, pengetahuan asosiasi antar jenis item ini dapat memberikan pengetahuan mengenai kombinasi item yang harus tercakup dalam promosi tersebut [4]. Dalam bidang usaha seperti swalayan, metode ini dikenal dengan istilah analisis keranjang pasar atau *market basket analysis*. *Market basket analysis* bisa menjadi tidak berguna ketika dihadapkan pada data yang kecil, namun akan menghasilkan informasi yang berharga ketika diperlakukan pada data transaksi yang besar seperti misalnya data transaksi sebuah *supermarket* [5].

Association rules mining memiliki beberapa algoritma yang sudah sering digunakan diantaranya algoritma Apriori, *FP-Growth*, dan *CT-PRO* yang dimana ketiga algoritma tersebut memiliki kekurangan dan kelebihan masing-masing [5]. *FP-Growth* adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. *FP-Growth* menggunakan pendekatan yang berbeda dari paradigma yang selama ini sering digunakan, yaitu paradigma apriori [6]. Algoritma *FP-Growth* merupakan pengembangan dari algoritma Apriori yang lebih efisien daripada algoritma Apriori. Algoritma ini hanya melakukan 2 kali proses *scanning database* untuk menentukan *frequent itemset* dan juga tidak

membutuhkan *generate candidate* seperti yang dilakukan pada algoritma Apriori. *FP-Growth* menggunakan konsep pembangunan *tree* disebut dengan *FP-Tree* dalam pencarian sehingga pemrosesanpun lebih cepat karena *frequent itemset* dapat langsung diekstrak dari hasil *tree* tersebut [7].

Dalam penelitian ini, akan dilakukan analisa terhadap data transaksi penjualan pada Gelael Candi Semarang menggunakan aturan asosiasi. Alogoritma yang digunakan adalah algoritma *FP-Growth* untuk menentukan himpunan yang sering muncul yang nantinya akan dijadikan sebagai acuan untuk merumuskan aturan-aturan asosiasi. Hasil aturan asosiasi tersebut dianalisa untuk menentukan barang apa saja yang memiliki keterkaitan yang dilihat dari segi kebiasaan konsumen membeli barang tersebut.

Berdasarkan uraian latar belakang diatas, maka peneliti memilih judul **“Metode Association Rule Dalam Menganalisa Pola Belanja Konsumen Pada Data Transaksi Penjualan Menggunakan Algoritma FP-Growth”**

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, permasalahan yang muncul adalah:

1. Bagaimana menganalisa pola belanja konsumen berdasarkan data transaksi penjualan swalayan sehingga dapat menampilkan suatu pengetahuan yang berguna bagi swalayan untuk meningkatkan pendapatan.
2. Bagaimana menerapkan metode asosiasi yang dilakukan menggunakan algoritma *FP-Growth* untuk menentukan keterkaitan antar barang sehingga dapat digunakan untuk strategi promosi produk.

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah :

1. Data yang gunakan adalah data transaksi penjualan yang diperoleh dari Swalayan Gelael Candi Semarang.
2. Metode yang digunakan adalah *association rule* dan algoritma yang dipakai adalah algoritma FP-Growth.
3. Nota transaksi yang hanya berisi 1 item tidak diikutsertakan.
4. Menggunakan aplikasi atau *tools* sebagai pengujian data.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Menggali pengetahuan dari analisa pada data transaksi penjualan untuk meningkatkan pendapatan.
2. Menentukan keterkaitan antar barang dengan metode asosiasi yang dilakukan menggunakan algoritma FP-Growth untuk strategi promosi produk.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

1. Manfaat bagi Penulis
 - a. Dapat menerapkan ilmu mengenai *Data Mining* yang telah diperoleh pada perkuliahan.
 - b. Dapat mengetahui lebih jauh mengenai penerapan metode Algoritma FP-Growth pada *Data Mining*.
 - c. Dapat memotivasi penulis untuk lebih banyak melakukan penelitian terkait dengan perkembangan ilmu pengetahuan dan teknologi.
2. Manfaat bagi Akademik
 - a. Dapat menambah koleksi jumlah penelitian terkait dengan perkembangan ilmu pengetahuan dan teknologi.

- b. Dapat meningkatkan minat penelitian di kalangan mahasiswa khususnya mengenai bidang ilmu *Data Mining*.

3. Manfaat bagi Perusahaan

- a. Membantu menganalisa produk yang sering dibeli secara bersamaan.
- b. Mengetahui produk yang paling sering dibeli oleh konsumen guna untuk meningkatkan persediaan.
- c. Membantu menentukan strategi promosi berdasarkan keterkaitan antar item yang sering dibeli oleh konsumen.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Dalam melaksanakan penelitian, peneliti mengambil beberapa referensi salah satunya yaitu referensi dari beberapa jurnal dengan tema dan metode yang berkaitan dengan penelitian.

2.1.1. Implementasi Algoritma FP-Growth Pada Aplikasi Prediksi Persediaan Sepeda Motor

Jurnal ini membahas tentang jumlah persediaan sepeda motor yang terlalu sedikit dan permintaan tidak dapat dipenuhi karena kekurangan persediaan, mengakibatkan konsumen merasa kecewa dan kemungkinan konsumen tidak akan kembali lagi, sehingga dapat menurunkan pendapatan perusahaan. Begitu juga jika persediaan terlalu banyak dan permintaan terlalu sedikit, akan menyebabkan kerugian dan perusahaan harus menyediakan tempat yang luas, dan harus menyediakan biaya tambahan seperti pemeliharaan dan biaya akuntansi.

Dalam jurnal ini dibahas suatu algoritma yaitu *FP-Growth*, dimana algoritma ini merupakan bagian dari teknik asosiasi dalam *Data Mining*. *FP-Growth* salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemsets*) dalam sekumpulan data. Dari hasil analisa, untuk mengimplementasikan algoritma *FP-Growth* pada aplikasi prediksi persediaan sepeda motor dapat dilihat dari banyaknya item yang terjual, lalu membuat rangkaian *tree* dengan *FP-Tree* untuk mengetahui banyaknya *frequent itemset* yang terjadi [8].

2.1.2.Implementasi *Data Mining* Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan

Jurnal ini membahas tentang ketersediaan barang di suatu Apotek dan jenis barang yang sering atau paling banyak dicari oleh konsumen yang menjadi prioritas utama Apotek yang harus distok lebih banyak untuk mengantisipasi kekosongan barang. Karena minimnya stok dapat berpengaruh pada pelayanan konsumen dan pendapatan Apotek. Sebagai salah satu tempat yang menyediakan alat-alat kesehatan, Apotek harus memperhatikan ketersediaan barang-barang yang dijual agar aktivitasnya dalam pelayanan terhadap konsumen berjalan dengan baik dan pendapatan yang diperoleh semakin meningkat.

Data Mining dapat diimplementasikan dengan menggunakan *database* penjualan alat-alat kesehatan karena dapat menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok jenis barang apa yang diperlukan kemudian. Hasil penerapan Algoritma Apriori pada teknik *Data Mining* sangat efisien dan dapat mempercepat proses pembentukan kecenderungan pola kombinasi itemset hasil penjualan alat-alat kesehatan di Apotek Kelambir-2 Medan, yaitu dengan support dan confidence tertinggi adalah Stick Asam Urat Stick Gula dan Stick Colestrol-Stick Gula [9].

2.1.3.Penerapan Metode Data Mining *Market Basket Analysis* Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan *Frequent Pattern Growth (FP-Growth)*

Jurnal ini membahas tentang eksperimen terhadap data transaksi penjualan produk buku di Percetakan PT. Gramedia dengan menggunakan sebuah metode *market basket analysis* untuk melihat hubungan asosiasi (kolerasi) antara sejumlah atribut penjualan. Algoritma yang digunakan adalah algoritma apriori dan algoritma *frequent pattern growth (FP-Growth)*. Kedua algoritma ini digunakan untuk membentuk *frequent*

itemsets yang dijadikan sebagai acuan dalam merumuskan aturan-aturan asosiasi yang dihasilkan oleh model *market basket analysis*.

Hasil yang didapat pada *market basket analysis* dengan menggunakan algoritma Apriori dan *FP-growth* terhadap atribut- atribut penjualan produk buku telah menghasilkan sejumlah aturan asosiasi yang berbeda antara algoritma yang satu dengan yang lainnya [6].

2.1.4. Association Rules Mining Untuk Penentuan Rekomendasi Promosi Produk

Jurnal ini membahas tentang perekomendasian promosi produk dengan memanfaatkan pendapat konsumen terhadap suatu barang untuk membantu pelanggan dalam memilih produk. Rekomendasi promosi produk memerlukan model rekomendasi yang tepat agar apa yang direkomendasikan sesuai dengan keinginan konsumen, serta mempermudah konsumen mengambil keputusan yang tepat dalam menentukan produk yang akan dibelinya. Hasil dari pembahasan jurnal ini menyimpulkan bahwa penerapan *Associastion Rule* dengan menggunakan algoritma *FP-Growth* dapat membantu penentuan rekomendasi promosi produk dengan tepat. Penerapan algoritma *FP-Growth* menghasilkan 3 *rule* ketika menggunakan *minimum support* 40% dan *minimum confidence* 80% [10].

Tabel 2.1 *State of the art* penelitian dengan metode *association rules*

No	Pengarang	Masalah	Metode	Hasil penelitian
1.	Ririanti (2014)	Untuk mengetahui persediaan sepeda motor maka dibutuhkan suatu aplikasi untuk memperoleh informasi mengenai jumlah sepeda motor.	Dalam penelitian menggunakan algoritma <i>FP-Growth</i>	Aplikasi untuk memprediksi persediaan sepeda motor.

No	Pengarang	Masalah	Metode	Hasil penelitian
2.	Kennedi Tampubolon, Hoga Saragih, Bobby Reza (2013)	Sebagai penyedia jasa, Apotek Kelambir-2 Medan masih melakukan pengecekan data barang secara manual sehingga tidak efektif dan efisien. Sehingga dibutuhkan analisa untuk mengetahui jumlah stok barang yang tersedia berdasarkan data transaksi penjualan.	Metode yang digunakan adalah <i>Association Rule</i> dengan menggunakan Algoritma Apriori	Menemukan kecenderungan pola kombinasi <i>itemsets</i> , membentuk pola kecenderungan kombinasi <i>itemsets</i> hasil penjualan alat-alat kesehatan.
3.	Goldi Gunadi dan Dana Indra Sensuse (2013)	Percetakan PT. Gramedia telah melayani sekian banyak transaksi pesanan cetakan. Setiap data transaksi tersebut disimpan di dalam suatu sistem basis data melalui aplikasi sistem informasi manajemen.	Metode yang digunakan adalah <i>Market Basket Analysis</i> dengan membandingkan 2 algoritma, yaitu Apriori dan <i>FP-Growth</i>	Menentukan algoritma yang paling baik untuk menganalisa data transaksi penjualan buku.

No	Pengarang	Masalah	Metode	Hasil penelitian
4.	Wiwit Agus Triyanto (2014)	Promosi Produk memerlukan model rekomendasi yang tepat agar apa yang direkomendasikan sesuai dengan keinginan pelanggan, serta mempermudah pelanggan mengambil keputusan yang tepat dalam menentukan produk yang akan dibelinya.	Metode yang digunakan adalah <i>Association Rule</i> dan Algoritma <i>FP-Growth</i>	Penentuan rekomendasi promosi produk yang tepat dan sesuai dengan keinginan pelanggan.

2.2 Landasan Teori

Peneliti meninjau beberapa buku dan jurnal sebagai landasan untuk menjelaskan *Data Mining*, diantaranya adalah sebagai berikut:

2.2.1 Definisi Data Mining

Banyak pendapat yang mendefinisikan *Data Mining*, secara umum *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 yang besar.

Menurut Turban dalam bukunya yang berjudul “*Decision Support System and Intelligent Systems*”, *Data Mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam basis data. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika,

kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar [11]. Menurut Gartner Group *data mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan menggunakan teknik pengenalan pola seperti statistik dan matematika [12].

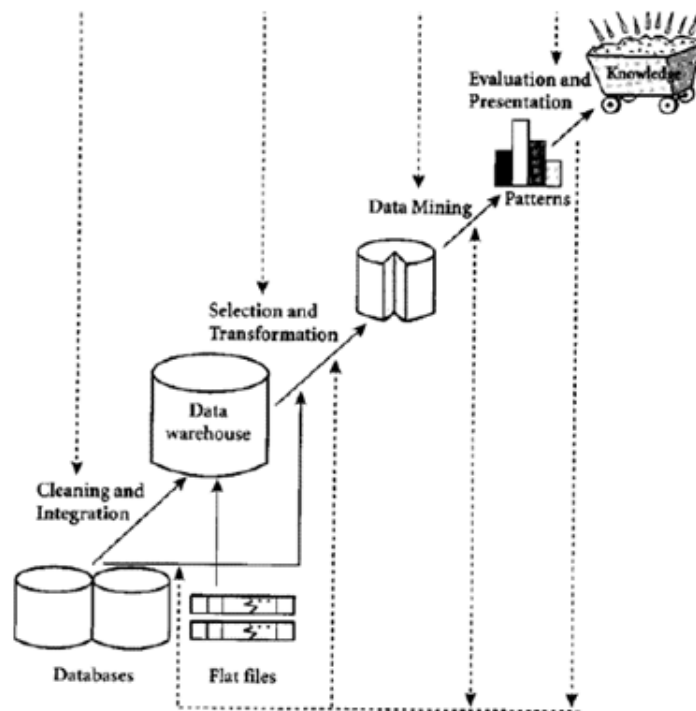
Kemajuan luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa faktor, antara lain [12]:

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam data *warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam database yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining.
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

Data mining, sering juga disebut *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Sehingga istilah *pattern recognition* jarang digunakan karena termasuk bagian dari data mining [13].

2.2.2 Tahap-Tahap Data Mining

Tahapan proses dalam penggunaan *Data Mining* yang merupakan proses *Knowledge Discovery in Database* (KDD) seperti yang diilustrasikan pada Gambar 2.1. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantara *knowledge base*.



Gambar 0.1 Tahap proses data mining

Tahapan-tahapan data mining [14]:

1. Pembersihan Data (*Data Cleaning*)

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi

performasi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (*Data Integration*)

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (*Data Selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data (*Data Transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses *Mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (*Pattern Evaluation*)

Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

7. Presentasi pengetahuan (*Knowledge Presentation*),

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

2.2.3 Teknik Data Mining

Data mining adalah proses ekstraksi data dengan jumlah besar untuk menggali pengetahuan yang selama ini belum diketahui dengan cara manual. Kata mining menunjukan bahwa proses data mining ini dilakukan dengan tujuan mendapatkan data berharga dari dalam database yang besar. Dari berbagai macam definisi data mining yang sangat luas, ada banyak teknik yang digunakan dalam data mining untuk mencari pengetahuan baru dalam

basis data. Berikut adalah teknik data mining dikelompokkan berdasarkan tugas yang dapat dilakukan, yaitu [12]:

1. Deskripsi

Tekadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapat dapat dipisahkan dalam tiga kategori, yaitu : pendapat tinggi, pendapat sedang, pendapat rendah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah

- Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori apa.

3. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable target estimasi lebih kearah numeric daripada kearah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Seelanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan dari sistolik dan nilai variabel prediksi dalam proses

pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

Contoh lain yaitu estimasi nilai indeks prestasi kumulatif mahasiswa pascasarjana dengan melihat nilai indeks prestasi mahasiswa tersebut pada saat mengikuti program sarjana.

4. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

Contoh prediksi dalam bisnis dan penelitian adalah :

- Prediksi harga beras dalam tiga bulan mendatang.
- Prediksi presentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

5. Pengklusteran

Pengklusteran merupakan pengumpulan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan record-record dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

Contoh pengklusteran dalam bisnis dan penelitian adalah :

- Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku financial dalam baik dan mencurigakan.
- Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis sering disebut analisis keranjang belanja.

Contoh asosiasi dalam bisnis dan penelitian adalah :

- Meneliti jumlah pelanggan dari perusahaan telekomunikasi selular yang diharapkan untuk memberikan respons positif terhadap penawaran upgrade layanan yang diberikan.
- Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

2.2.4 Market Basket Analysis

Market basket analysis adalah suatu metodologi untuk melakukan analisis *buying habit* konsumen dengan menemukan asosiasi antar beberapa item yang berbeda, yang diletakkan konsumen dalam *shopping basket* yang dibeli pada suatu transaksi tertentu. Tujuan dari *market basket* adalah untuk mengetahui produk-produk mana yang mungkin akan dibeli secara bersamaan. Analisis data transaksi dapat menghasilkan pola pembelian produk yang sering terjadi. Teknik ini telah banyak digunakan oleh toko grosir maupun retail [11].

Market Basket Analysis memanfaatkan data transaksi penjualan untuk dianalisis kemudian menemukan pola item-item yang secara bersamaan dalam suatu transaksi. Salah satu manfaat dari *Market Basket Analysis* merancang strategi penjualan atau pemasaran dengan memanfaatkan data penjualan yang ada di perusahaan yaitu :

1. Dengan mengubah tata letak toko, menempatkan item-item barang secara berdekatan yang sering di beli secara bersamaan oleh konsumen.
2. Memberikan diskon kepada item barang yang jarang di beli dan mahal.

2.2.5 Association Rule

Analisa asosiasi atau *association rule* mining adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item. Contoh aturan asosiasi dari analisis pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut, pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu. Karena analisis asosiasi menjadi terkenal karena aplikasinya untuk menganalisis isi keranjang belanja dipasar swalayan, analisis asosiasi juga sering disebut *Market Basket Analysis* [15].

Menurut Erwin [16], ada dua ukuran parameter dalam analisa asosiasi yaitu:

1. *Support* adalah pengukuran untuk menunjukkan seberapa besar tingkat dominasi suatu item dari keseluruhan transaksi.
2. *Confidence* adalah pengukuran untuk menunjukkan hubungan antar dua item berdasarkan suatu kondisi tertentu.

Kedua ukuran ini nantinya berguna dalam menentukan kekuatan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua parameter yang ditentukan oleh pengguna. Bila suatu pola memenuhi kedua nilai minimum parameter yang sudah ditentukan maka sebuah pola dapat disebut *interesting rule* atau *strong rule*. *Association Rule* biasanya dinyatakan dalam bentuk : {roti,mentega} \rightarrow {susu} (*support* = 40%, *confidence* = 50%) Yang artinya : "50% dari transaksi di *database* yang memuat *item* roti dan mentega juga memuat *item* susu. Sedangkan 40% dari seluruh transaksi yang ada di *database* memuat ketiga *item* tersebut. Dapat juga diartikan : "Seorang konsumen yang membeli roti dan mentega punya kemungkinan 50% untuk juga membeli susu. Aturan tersebut cukup akurat karena mewakili 40% dari catatan transaksi yang ada".

Metodologi dasar *Association Rule* terbagi menjadi dua tahap [15], yaitu :

a. Analisa Pola Frekuensi Tinggi

Tahap ini mencari pola *item* yang memenuhi syarat minimum dari nilai *support* dalam *database* . Nilai *support* sebuah *item* diperoleh dengan rumus berikut :

$$Support(A) = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total Transaksi}} \quad (1)$$

Pada rumus 1 menjelaskan bahwa nilai *support* diperoleh dengan cara mencari jumlah transaksi yang mengandung nilai A (satu *item*) dibagi dengan jumlah keseluruhan transaksi. Sedangkan nilai *support* dari 2 *item* diperoleh dari rumus berikut :

$$Support(A, B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \quad (2)$$

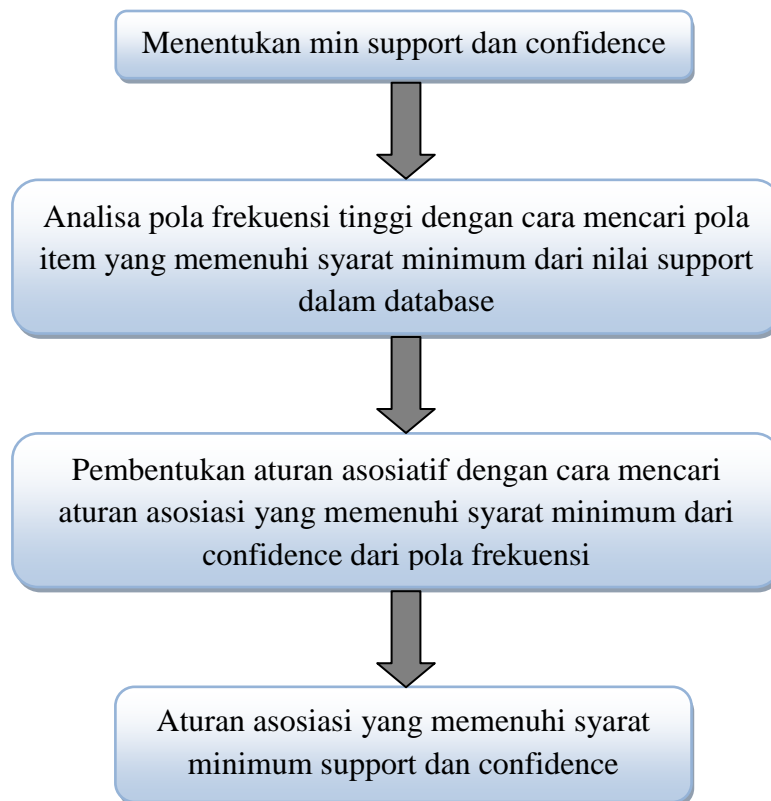
Pada rumus 2 menjelaskan bahwa nilai *support* diperoleh dengan cara mencari jumlah transaksi yang mengandung nilai A dan B (*item* pertama bersamaan dengan *item* yang lain) dibagi dengan jumlah keseluruhan transaksi.

b. Pembentukan Aturan Asosiatif

Setelah semua pola frekuensi tinggi ditemukan, kemudian mencari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan assosiatif A->B dari *support* pola frekuensi tinggi A dan B, menggunakan rumus berikut :

$$\begin{aligned} \text{Confidence} &= P(B|A) \\ &= \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \quad (3) \end{aligned}$$

Pada rumus 3 menjelaskan bahwa nilai *confidence* diperoleh dengan cara mencari jumlah transaksi yang mengandung nilai A dan B (*item* pertama bersamaan dengan *item* yang lain) dibagi dengan jumlah transaksi yang mengandung A (*item* pertama). Secara garis besar, metodologi *Association Rule* dapat dilihat pada gambar 2.2 dibawah ini.



Gambar 2.2 Metodologi *Association Rule*

2.2.6 Algoritma FP-Growth

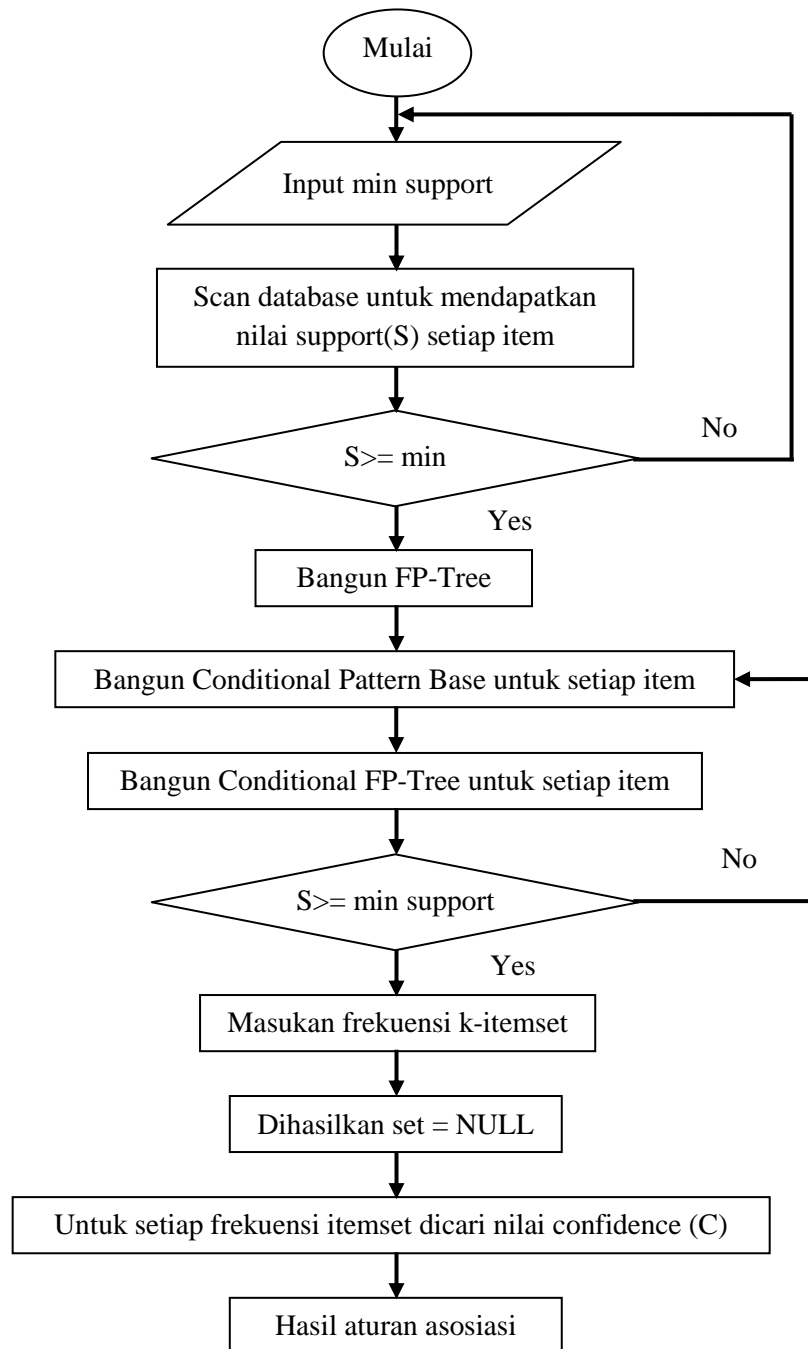
Menurut Yuan [17], algoritma *Frequent Pattern Growth (FP-Growth)* merupakan pengembangan dari algoritma Apriori. Algoritma ini memperbaiki kekurangan dari algoritma Apriori. Algoritma ini merupakan salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data.

Pada algoritma *FP-Growth generate candidate* tidak dilakukan karena *FP-Growth* menggunakan konsep *tree* dalam pencarian *frequent itemset*. Karakteristik algoritma *FP-Growth* adalah penggunaan struktur *tree* yang disebut dengan *FP-Tree* [16]. Dengan menggunakan *FP-Tree* dapat mengetahui *frequent itemset* dari *FP-Tree* tersebut. Menurut [6], proses algoritma *FP-Growth* adalah:

2.2.6.1. Pembuatan *FP-Tree*

FP-Tree adalah salah satu struktur data yang berbentuk *tree*. *FP-Tree* dibangun dengan memetakan setiap data transaksi kedalam setiap lintasan tertentu dalam *FP-Tree*. Karena dalam setiap transaksi yang dipetakan, mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data *FP-tree* semakin efektif. *FP-tree* adalah sebuah pohon dengan definisi sebagai berikut [6]:

- a. *FP-tree* dibentuk oleh sebuah akar yang diberi label null, sekumpulan *sub-tree* yang beranggotakan item-item tertentu, dan sebuah tabel *frequent header*.
- b. Setiap simpul dalam *FP-tree* mengandung tiga informasi penting, yaitu label item, menginformasikan jenis item yang direpresentasikan simpul tersebut, *support count*, merepresentasikan jumlah lintasan transaksi yang melalui simpul tersebut, dan *pointer* penghubung yang menghubungkan simpul-simpul dengan label item sama antar-lintasan, ditandai dengan garis panah putus-putus.



Gambar 2.3 Diagram alir algoritma *FP-Growth*

2.2.6.2. Penerapan Algoritma FP-Growth

Setelah sekumpulan data transaksi dibangun dalam struktur *FP-Tree*, algoritma untuk mencari *frequent itemset* yang signifikan akan diterapkan. Algoritma *FP-Growth* terdiri dari tiga langkah utama, yaitu [6]:

a. Tahap Pembangkitan *Conditional Pattern Base*

Conditional Pattern Base merupakan *subdatabase* yang berisi *prefix path* (lintasan *prefix*) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui *FP-tree* yang telah dibangun sebelumnya.

b. Tahap Pembangkitan *Conditional FP-Tree*

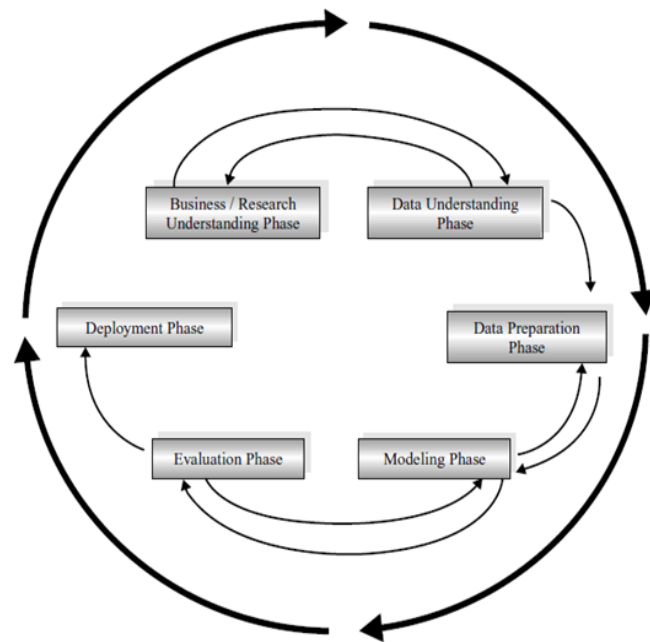
Pada tahap ini, *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah *support count* lebih besar sama dengan minimum *support count* akan dibangkitkan dengan *conditional FP-Tree*.

c. 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.

2.2.7 CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) yang dikembangkan pada tahun 1996 oleh analis dari beberapa industri seperti Daimler Chrysler, SPSS, dan NCR. CRISP-DM menyediakan standar proses *data mining* sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian. Sebuah proyek *data mining* dalam CRISP-DM memiliki siklus hidup yang terbagi dalam lima fase (Larose, 2005), yaitu [11]:



Gambar 2.4 Proses Data Mining Menurut CRISP-DM

Dari Gambar 2.4 Enam fase CRISP-DM, diantaranya:

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
 - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.
2. Fase Pemahaman Data (*Data Understanding Phase*)
 - a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.

- d. Jika diinginkan, pilih sebagian kecil group data yang mungkin mengandung pola dari permasalahan.
3. Fase Pengolahan Data (*Data Preparation Pharse*)
 - a. Siapkan dari data awal, kumpulkan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
 - b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.
 - c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
 - d. Siapkan data awal sehingga siap untuk perangkat pemodelan.
4. Fase Pemodelan (*Modeling Pharse*)
 - a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
 - b. Kalibrasi aturan model untuk mengoptimalkan hasil.
 - c. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *data mining* yang sama.
 - d. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.
5. Fase Evaluasi (*Evaluation Pharse*)
 - a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektifitas sebelum disebarkan untuk digunakan.
 - b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
 - c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak ditangani dengan baik.
 - d. Mengambil keputusan berkaitan dengan penggunaan hasil dari *data mining*.
6. Fase Penyebaran (*Deployment Pharse*)
 - a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.

- b. Contoh sederhana penyebaran : pembuatan laporan.
- c. Contoh kompleks penyebaran : penerapan proses *data mining* secara paralel pada departemen lain.

2.3 Kerangka Pemikiran

Masalah		
Bagaimana menerapkan Algoritma FP-Growth untuk menentukan strategi promosi pada Swalayan Gelael Candi.		
Tujuan		
Menerapkan Algoritma FP-Growth untuk menentukan strategi promosi pada Swalayan Gelael Candi		
Eksperimen/Analisa*		
Aplikasi	Data	Metode
Java (NetBeans)	Data transaksi penjualan pada Swalayan Gelael Candi Semarang	<i>Metode Assosiation rules dan Algoritma Frequent Pattern Growth (FP-Growth)</i>
Hasil		
Sebagai bahan masukan untuk mendukung pengambilan keputusan strategi promosi pada Swalayan Gelael Candi, memberikan informasi tentang pembelian produk yang sering dibeli oleh konsumen secara bersamaan, mengetahui bagaimana memanfaatkan data penjualan yang ada sehingga menjadi pengetahuan baru yang dapat digunakan untuk mencari solusi dalam meningkatkan penjualan.		

BAB III

METODE PENELITIAN

3.1 Instrumen Penelitian

Sebagai instrumen penelitian, akan dilakukan penelitian untuk melakukan implementasi algoritma FP-Growth untuk tujuan dari penelitian. Implementasi menggunakan Matlab untuk mengetahui keterkaitan antar item yang ada pada data transaksi.

3.2 Objek Penelitian

Objek penelitian dilakukan pada sebuah swalayan yang didirikan secara perorangan, swalayan tersebut adalah Gelael Candi Semarang.

3.3 Sumber Data

Sumber data yang digunakan dalam penelitian ini yaitu

a. Data Primer

Data primer yaitu data yang diperoleh secara langsung dari sumber data yang berhubungan dengan penelitian yang dilakukan. Dalam hal ini data yang digunakan adalah data transaksi penjualan selama satu bulan. Data yang berjumlah besar ini nantinya akan dilakukan data mining untuk menghasilkan rule belanja konsumen dan dapat menentukan produk mana yang paling sering dibeli secara bersamaan.

b. Data Sekunder

Data yang diperoleh dari data penelitian dalam bentuk yang sudah jadi yang bersifat informasi dan kutipan, baik dari internet maupun dari literatur yang berkaitan dengan permasalahan yang dihadapi dan diteliti.

3.4 Metode Pengumpulan Data

1. Studi Pustaka

Studi pustaka yaitu menerapkan metode pengumpulan data yang dilakukan dengan mencari, membaca, dan mengumpulkan dokumen-dokumen sebagai referensi seperti buku, jurnal, artikel, dan literatur lain yang berhubungan dengan topic yang digunakan oleh penulis untuk mendapatkan tambahan informasi tentang metode yang digunakan penulis dalam penelitian.

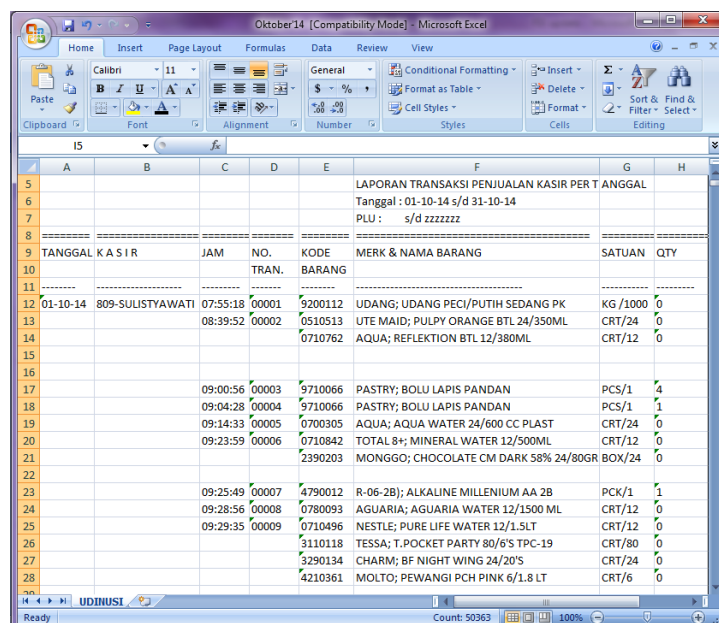
2. Wawancara

Wawancara dilakukan untuk memperoleh data tentang produk yang dijual pada swalayan Gelael Candi Semarang. Wawancara dilakukan dengan Manager Gelael Candi Semarang.

3.5 Tahap Penelitian

3.5.1. Pengumpulan Data

Penulis mengumpulkan data dengan meminta data transaksi selama 1 (satu) bulan yaitu bulan Oktober pada Swalayan Gelael Candi Semarang. Bentuk datanya yaitu dalam bentuk file *microsoft office excel*.



	A	B	C	D	E	F	G	H
5						LAPORAN TRANSAKSI PENJUALAN KASIR PER TANGGAL		
6						Tanggal : 01-10-14 s/d 31-10-14		
7						PLU : s/d zzzzzzz		
8								
9	TANGGAL	KASIR	JAM	NO. TRAN.	KODE BARANG	MERK & NAMA BARANG	SATUAN	QTY
10								
11								
12	01-10-14	809-SULISTYAWATI	07:55:18	00001	9200112	UDANG; UDANG PEPI/PUTIH SEDANG PK	KG/1000	0
13			08:39:52	00002	0510513	UTE MAID; PULPY ORANGE BTL 24/350ML	CRT/24	0
14					0710762	AQUA; REFLEKSION BTL 12/380ML	CRT/12	0
15								
16			09:00:56	00003	9710066	PASTRY; BOLU LAPIS PANDAN	PCS/1	4
17			09:04:28	00004	9710066	PASTRY; BOLU LAPIS PANDAN	PCS/1	1
18			09:14:33	00005	0700305	AQUA; AQUA WATER 24/600 CC PLAST	CRT/24	0
19			09:23:59	00006	0710842	TOTAL 8+; MINERAL WATER 12/500ML	CRT/12	0
20					2390203	MONGGO; CHOCOLATE CM DARK 58% 24/80GR BOX/24		0
21								
22			09:25:49	00007	4790012	R-06-2B); ALKALINE MILLENIUM AA 2B	PCK/1	1
23			09:28:56	00008	0780093	AGUARIA; AGUARIA WATER 12/1500 ML	CRT/12	0
24			09:29:35	00009	0710496	NESTLE; PURE LIFE WATER 12/1.5LT	CRT/12	0
25					3110118	TESSA; T-POCKET PARTY 80/6'S TPC-19	CRT/80	0
26					3290134	CHARM; BF NIGHT WING 24/20'S	CRT/24	0
27					4210361	MOLTO; PEWANGI PCH PINK 6/1.8 LT	CRT/6	0
28								

Gambar 3.1 Sample dataset transaksi penjualan

Pada gambar 3.1 tabel data transaksi terdiri dari 8 atribut, yaitu tanggal transaksi, kasir, jam, nomor transaksi, kode barang, merk dan nama barang, satuan, dan *quantity*. Akan tetapi pada *association rule* tidak semua atribut diikuti sertakan dalam pengolahannya. Berikut adalah penjelasan tentang atribut yang ada pada dataset:

Tabel 3.1 Penjelasan atribut dalam dataset

ATRIBUT	KETERANGAN
Tanggal Transaksi	Tanggal saat transaksi berlangsung
Kasir	Nama kasir
Jam	Waktu saat transaksi berlangsung
Nomor Transaksi	Nomer transaksi penjualan
Kode Barang	Kode barang yang dibeli
Merk dan Nama Barang	Nama barang yang dibeli
Satuan	Satuan barang
Quantity	Jumlah barang yang dibeli

3.5.2. Penyeleksian Data

Dari data yang dikumpulkan, dilakukan penyeleksian dengan memilih dan memisahkan data transaksi berdasarkan kategori item yang ditentukan. Dari data transaksi yang ada, tidak semua ikut dalam proses pengolahan. Pada penelitian ini yang digunakan hanya data transaksi penjualan yang mempunyai lebih dari 1 jenis produk dan tidak termasuk dalam 1 departemen pengelompokan barang.

3.5.3. Preprocessing/Cleaning

Pada tahap ini dilakukan proses cleaning pada data transaksi penjualan yang menjadi fokus data mining. Dilakukan proses penyeleksian dengan memilih atribut yang akan digunakan dalam proses data mining. Atribut yang dipakai yaitu No. Transaksi, Kode Barang, Nama Barang, selain atribut tersebut akan dihapus.

Tabel 3.2 Sample data yang akan diolah

Nomor Transaksi	Kode Barang
00001	9200112
00002	0510513
	0710762
00003	9710066
00004	9710066
00005	0700305
00006	0710842
	2390203
00007	4790012
00008	0780093
00009	0710496
	3110118
	3290134
	4210361
00010	5601189
	9700171
00011	4210113
00012	0700305
	0801083
	1510041
	1510371
	2110795
	2190100
	3710178
	5200362
	8200018
00013	8700057
	9190038

	9190038
00014	0700294
00015	8200014
	8300040
00016	0780093
00017	3010015
	7001417
00018	0610120
	1500022
	1810031
	2500053
	2900005
	3310338
	4010011
	4010483
	4210634
	6490075
	8510011
	8780031
00019	0700294
	0710632
00020	0710568
	0780093

Tabel 3.3 Data transaksi dengan nama barang

Nomor Transaksi	Kode Barang	Nama Barang
00001	9200112	UDANG; UDANG PECI/PUTIH SEDANG PK
00002	0510513	UTE MAID; PULPY ORANGE BTL 24/350ML
	0710762	AQUA; REFLEKTION BTL 12/380ML
00003	9710066	PASTRY; BOLU LAPIS PANDAN
00004	9710066	PASTRY; BOLU LAPIS PANDAN
00005	0700305	AQUA; AQUA WATER 24/600 CC PLAST
00006	0710842	TOTAL 8+; MINERAL WATER 12/500ML
	2390203	MONGGO; CHOCOLATE CM DARK 58% 24/80GR
00007	4790012	R-06-2B); ALKALINE MILLENIUM AA 2B
00008	0780093	AGUARIA; AGUARIA WATER 12/1500 ML
00009	0710496	NESTLE; PURE LIFE WATER 12/1.5LT
	3110118	TESSA; T.POCKET PARTY 80/6'S TPC-19
	3290134	CHARM; BF NIGHT WING 24/20'S
	4210361	MOLTO; PEWANGI PCH PINK 6/1.8 LT
00010	5601189	DEN FARM; SHOESTRING 12/1 KG
	9700171	TI TAWAR; CASINO BREAD (KULIT)
00011	4210113	EAR(REF); PEMBERSIH KACA POUCH BR 12/4
00012	0700305	AQUA; AQUA WATER 24/600 CC PLAST
	0801083	BINTANG; BEER CAN 24/330 ML
	1510041	APAL API; KOPI SPC MIX BOX 24/5/25GR

	1510371	PI LUWAK; KOPI BUBUK ARABIKA 24/165GR
	2110795	HAPPYTOS; TORTILLA R.CORN BIRU 32/55GR
	2190100	CHITATO; PICNIC SIZE R.AYAM BUMBU 30/7
	3710178	GO FRESH; MINI PACK PEPPERMINT 10/24'S
	5200362	NG GARAM; FILTER SURYA 10/12'S
	8200018	PEPAYA; PEPAYA
00013	8700057	DILMAH; P.CAMOMILE FLOW.TEA 12/37,5GR
	9190038	E & VIRE; WHIPPING CREAM 1 L
	9190038	E & VIRE; WHIPPING CREAM 1 L
00014	0700294	AQUA; AQUA WATER 12/1,5 LTR PLAS
00015	8200014	SALAK; SALAK PONDOH SUPER A
	8300040	ANGGUR; ANGGUR MERAH (RED GLOBE) IMPOR
00016	0780093	AGUARIA; AGUARIA WATER 12/1500 ML
00017	3010015	KLIN PAK; KANTONG SAMPAH BESAR 24/8'S
	7001417	WEINA; KUAS MARGARIN 2'S
00018	0610120	UTRISARI; AMERICAN S.ORANGE 72/10/12 G
	1500022	INDOCAFE; COFFEEMIX 12/30'S
	1810031	SANIA; MINYAK GORENG POUCH 12/1 LTR
	2500053	INDOMIE; MIE GORENG SPESIAL 40/84 GR

	2900005	TELUR; TELUR AYAM NEGERI KILO
	3310338	LUX; B.WASH PCH PINK 24/450ML
	4010011	SUNLIGHT; LIQUID REF LIME 12/800ML
	4010483	ATTACK; DETERGENT+SOFTENNER 8/1.2KG
	4210634	TO ULTRA; SOFTENER PCH PURE 12/900ML
	6490075	MORINAGA; CHIL KID BOX 200 G MADU
	8510011	GULAKU; GULA TEBU PREMIUM (PTH) 24/1KG
	8780031	TONG TJI; TEH CELUP JASMINE 40/25/50 G
00019	0700294	AQUA; AQUA WATER 12/1,5 LTR PLAS
	0710632	RI SWEAT; ISOTONIC DRINK PET 6/2LTR
00020	0710568	RI SWEAT; ISOTONIC DRINK PET 24/350ML
	0780093	AGUARIA; AGUARIA WATER 12/1500 ML

3.5.4.Transformasi Data

Data transaksi yang telah melalui proses seleksi dan preprocessing / cleaning ditransformasikan berdasarkan kode setiap departemen dan disimpan dalam bentuk .txt.

Tabel 3.4 Dataset hasil transformasi

TRANSAKSI ID	ITEM
00001	5 6 7 21 75 87
00002	7 14 21 76
00003	2 7 20 21
00004	5 20 40 47 81
00005	7 75
00006	15 75 76 81
00007	21 76 87
00008	19 54
00009	11 22
00010	5 21 81 87
00011	5 20 23 81
00012	81 87 88
00013	28 64
00014	7 21 75
00015	76 78 80 81
00016	17 31 56
00017	5 7 59
00018	7 20 40
00019	21 22 76 88
00020	6 31
...dst	...

3.5.5.Data Mining

Pada tahap ini dipusatkan untuk mendapatkan pola dari ekstraksi data transaksi yang sudah ditransformasi dengan menerapkan algoritma FP-Growth. Proses data mining dilakukan dengan membentuk sebuah *frequent item set* dengan menentukan nilai *support* dan mendapatkan nilai kepastian dengan menentukan nilai *confidence*.

3.5.6.Interpretation/Evaluation

Pola-pola yang telah diidentifikasi oleh sistem kemudian diterjemahkan atau diinterpretasikan ke dalam bentuk yang bisa dimengerti untuk membantu dalam perencanaan strategi bisnis bagi Swalayan Gelael Candi.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Sumber data yang digunakan berasal dari data transaksi di Swalayan Gelael Candi Semarang pada bulan Oktober 2014. Data transaksi tersebut akan diolah untuk mendapatkan pengetahuan baru berupa pola asosiasi antar barang yang dapat digunakan untuk strategi bisnis.

4.2 Pengumpulan Data

Item yang berada di Swalayan Gelael Candi sangatlah beraneka ragam, sehingga perlu adanya pengelompokan item. Dalam penelitian ini pengelompokan dibagi menjadi 89 departemen. Adapun 89 departemen pengelompokan tersebut adalah :

- | | |
|------------------------------|----------------------------|
| 1. Buah Dalam Kaleng | 16. Biji-Bijian |
| 2. Sayuran Dalam Kaleng | 17. Minyak Goreng |
| 3. Daging Dalam Kaleng | 18. Preecook |
| 4. Seafood Dalam Kaleng | 19. Wafer dan Biscuits |
| 5. Juice Buah | 20. Salties |
| 6. Syrup | 21. Cookies Kering |
| 7. Soft Drink | 22. Coklat Bar |
| 8. Susu | 23. Permen dan Jelly |
| 9. Jam, Meisis dan Madu | 24. Mie Instant |
| 10. Sauce, Kecap dan Sambal | 25. Ikan Asin |
| 11. Herb Biji | 26. Daging Kering |
| 12. Soups | 27. Telor |
| 13. Cake Make and Decoration | 28. Napkins dan Pembungkus |
| 14. Kopi | 29. Tissue |
| 15. Cereal | 30. Kapas dan Pembalut |

31. Sabun Mandi
32. Kosmetik
33. Parfume dan Deodorant
34. Shampoo dan Conditioner
35. Pasta Gigi
36. Hair Brush
37. Shaving dan Foam
38. Sabun Cuci
39. Desinfektant
40. Pemelihara dan Pembersih
41. Air Freshener dan Kamper
42. Aksesoris Mobil dan
Peralatan Tukang
43. Keperluan Hewan
44. Keperluan Tumbuhan
45. Batu Baterai
46. Kue Tradisional
47. Obat-obatan dan Vitamin
48. Perlengkapan tobacco
49. Tabacco
50. Wine dan Spirits
51. Sayur Segar
52. Beef / Chicken Sausages &
Burger
53. Sayuran Frozen
54. Margarin dan Mentega
55. Ice Cream
56. Japan dan Korean Food
57. Diabetic
58. Food Preserved Fruits &
Vegetables
59. Baby Food
60. Hiasan dan Gambar
61. Parcel dan Produk Branded
62. Baby Supply Non Food
63. Baby Care
64. Barang RT Non Electric
65. Glass Ware
66. Keramik Ware
67. Plastik/Melamine Ware
68. Barang RT Electric
69. Sikat Gigi
70. Talc Powder
71. Hair Care
72. Stainless dan Alumunium
73. Soft Goods
74. Pork Meat
75. Buah Segar Lokal
76. Buah Segar Import
77. Gula
78. Teh
79. Minuman Berenergi
80. Dressing
81. Susu Segar dan Yoghurt
82. Keju
83. Seafood Segar
84. Ayam Segar
85. Dog Food
86. Daging Segar
87. Bakery & Pastry
88. Kue Basah
89. Wet Market

4.3 Preprocessing/Cleaning

Data transaksi penjualan terdapat banyak atribut, tidak semua atribut-atribut tersebut diperlukan dalam pengolahan data mining. Maka dari itu diperlukan pembersihan/cleaning yang bertujuan untuk memilih atribut data yang menjadi fokus sesuai dengan algoritma yang digunakan dan menghilangkan atau menghapus atribut-atribut yang tidak diperlukan. Beberapa atribut yang terdapat dalam data transaksi penjualan diantaranya :

1. Tanggal transaksi
2. Kasir
3. Jam
4. Nomor transaksi
5. Kode barang
6. Nama dan merk barang
7. Satuan
8. Quantity

4.4 Transformasi Data

Tahap transformasi data perlu dilakukan, karena dalam proses *mining* secara komputerisasi diperlukan bentuk data yang bisa diintegrasikan dengan aplikasi yang akan digunakan. Data transaksi yang telah melalui tahap penyeleksian dan preprocessing/cleaning ditransformasikan dalam bentuk kode angka sesuai dengan departemen. Proses transformasinya adalah nama kategori *item* diubah menjadi angka sesuai dengan no. kategori *item* seperti yang sudah dijelaskan diatas.

4.5 Data Mining

Pada tahap ini, *FP-Tree* dibuat untuk digunakan bersamaan dengan algoritma *FP-Growth* untuk menentukan *frequent itemsets*, Data yang akan diolah sebagai contoh yaitu data sebanyak 20 transaksi seperti terlihat pada tabel 4.1. Penulis memberikan batasan nilai minimum *support* 0.15 atau sama dengan 15 %.

dan *confidence* 0.60 atau sama dengan 60%, Dalam membangun *FP-Tree* diperlukan dua kali penelusuran *database*.

Tabel 4.1 Contoh 20 data transaksi

ID	Item
1	{5,6,7,21,75,87}
2	{7,14,21,76}
3	{2,7,20,21}
4	{5,20,40,47,81}
5	{7,75}
6	{15,75,76,81}
7	{21,76,87}
8	{19,54}
9	{11,22}
10	{5,21,81,87}
11	{5,20,23,81}
12	{81,87,88}
13	{28,64}
14	{7,21,75}
15	{76,78,80,81}
16	{17,31,56}
17	{5,7,59}
18	{7,20,40}
19	{21,22,76,88}
20	{6,31}

Penelusuran *database* yang pertama digunakan untuk menghitung nilai *support* masing-masing *item* dan memilih *item* yang memenuhi nilai minimum *support*. Hasil dari proses penelusuran *database* yang pertama adalah diketahuinya jumlah frekuensi kemunculan tiap *item* yang ada didalam *database*

dan mengurutkannya berdasarkan jumlah frekuensi kemunculan terbesar. Hasilnya dapat dilihat pada tabel 4.2. Dari hasil tersebut diperoleh *Itemset* yang memiliki frekuensi diatas minimum *support count* $\geq 0,15$ yaitu 7,21,81,5,76,20,75 dan 87 yang kemudian diberi nama *Frequent List* seperti terlihat pada tabel 4.3 kedelapan *item* ini akan berpengaruh saat pembuatan *FP-Tree*. Item yang frekuensi kemunculannya dibawah minimum *support count* akan dihilangkan.

Tabel 4.2 Frequent itemset

<i>Item</i>	Frekuensi
7	7
21	7
81	6
5	5
76	5
20	5
75	4
87	4
6	2
22	2
31	2
40	2
88	2
2	1
11	1
14	1
15	1
17	1
19	1
23	1

28	1
47	1
54	1
56	1
59	1
64	1
78	1
80	1

Tabel 4.3 Frequent list dengan min *support* $\geq 0,15$

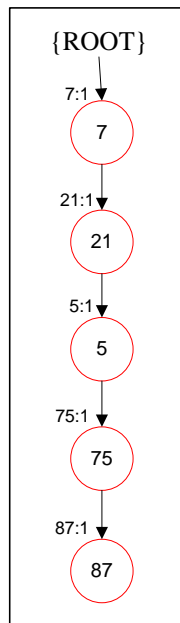
<i>Item</i>	Frekuensi
7	7
21	7
81	6
5	5
76	5
20	5
75	4
87	4

Setelah di peroleh Frequent List, kemudian item dirubah sesuai Frequent List dengan menghilangkan item yang tidak memenuhi minimum support yang telah ditentukan seperti item 6, 22, 31, 40, 88, 2, 11, 14, 15, 17, 19, 23, 28, 47, 54, 56, 59, 64, 78 dan 80. Urutkan data transaksi dari item yang memiliki nilai support terbesar hingga terkecil seperti pada table 4.4.

Tabel 4.4 Transaksi yang sudah diurutkan berdasarkan F-List

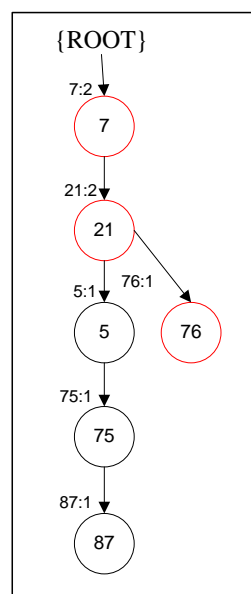
TID	Item
1	{7,21,5,75,87}
2	{7,21,76}
3	{7,21,20}
4	{81,5,20}
5	{7,75}
6	{81,76,75}
7	{21,76,87}
8	0
9	0
10	{21,81,5,87}
11	{81,5,20}
12	{81,87}
13	0
14	{7,21,75}
15	{81,76}
16	0
17	{7,5}
18	{7,20}
19	{21,76}
20	0

Setelah *item* disusun ulang berdasarkan *F-list*, dilakukan penelusuran *database* yang kedua yaitu membaca tiap transaksi diawali dengan membaca TID 1 untuk membuat *FP-Tree*. TID 1 {7,21,5,75,87} akan membuat simpul 5 sehingga terbentuk lintasan seperti gambar4.1 dengan *support count* awal bernilai satu.



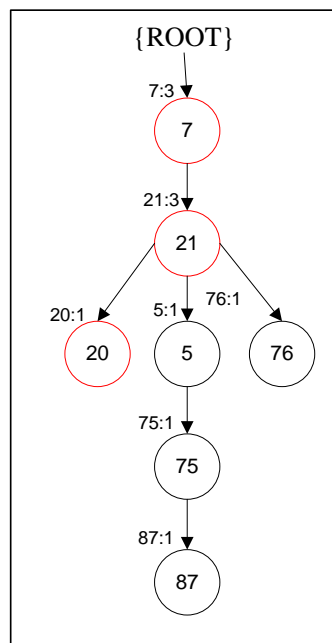
Gambar 0.1 Hasil pembentukan TID 1

Setelah pembacaan TID 1, selanjutnya membaca TID 2 yaitu {7,21,76}. Karena mempunyai *prefix* yang sama dengan lintasan TID 1 yaitu item 7 dan 21 maka lintasan yang terbentuk pada TID 2 bisa dipadatkan pada lintasan TID 1. Setelah itu tambahkan *support count* pada item 7 dan 21 menjadi dua karena telah dilewati sebanyak dua kali dan membentuk lintasan baru untuk item 76 dengan *support count* bernilai satu.



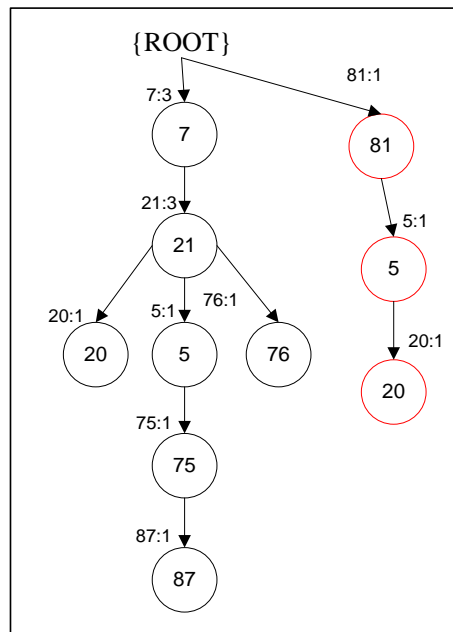
Gambar 0.2 Hasil pembentukan TID 2

Setelah pembacaan TID 2, selanjutnya membaca TID 3 yaitu {7,21,20}. Karena mempunyai *prefix* yang sama dengan lintasan TID 1 yaitu item 7 dan 21 maka lintasan yang terbentuk pada TID 3 bisa dipadatkan pada lintasan TID 1. Setelah itu tambahkan *support count* pada item 7 dan 21 menjadi tiga karena telah dilewati sebanyak tiga kali dan membentuk lintasan baru untuk item 20 dengan *support count* bernilai satu.



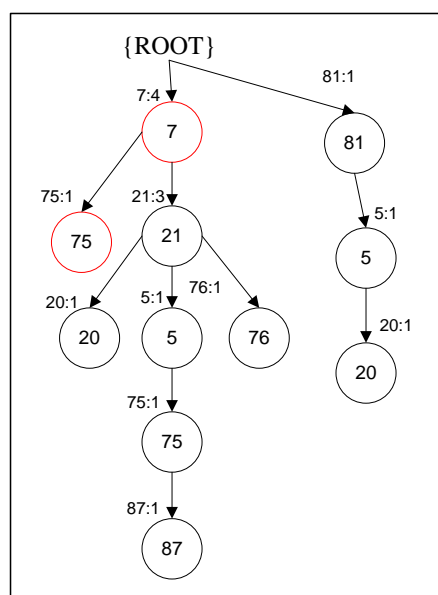
Gambar 0.3 Hasil pembentukan TID 3

Setelah pembacaan TID 3, selanjutnya adalah membaca TID 4 yaitu {81,5,20}. Karena prefix lintasannya tidak sama dengan lintasan sebelumnya, maka TID 4 dibuat lintasan baru.



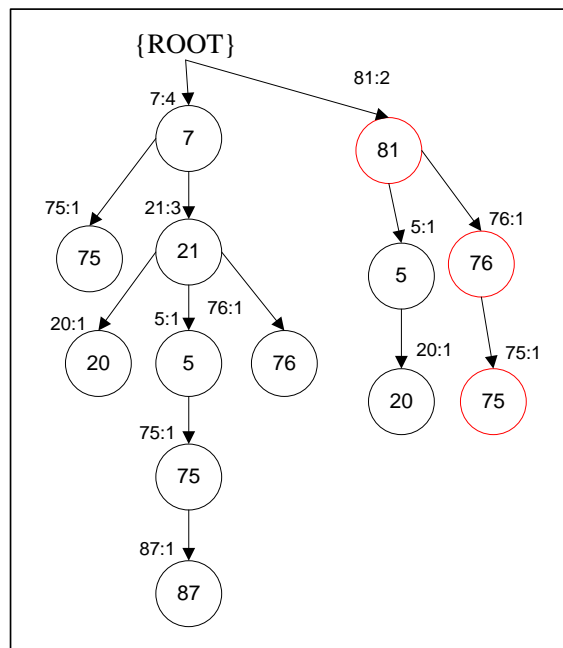
Gambar 0.4 Hasil pembentukan TID 4

Setelah pembacaan TID 4, selanjutnya adalah membaca TID 5 yaitu {7,75}. Karena mempunyai *prefix* yang sama dengan lintasan TID 1 yaitu item 7 maka lintasan yang terbentuk pada TID 5 bisa dipadatkan pada lintasan TID 1. Setelah itu tambahkan *support count* pada item 7 menjadi empat karena telah dilewati sebanyak empat kali dan membentuk lintasan baru untuk item 75 dengan *support count* bernilai satu.



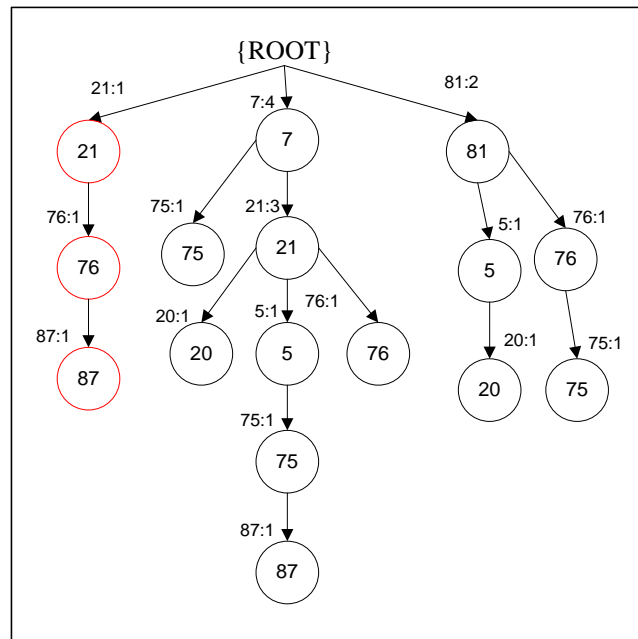
Gambar 0.5 Hasil pembentukan TID 5

Setelah pembacaan TID 5, selanjutnya adalah membaca TID 6 yaitu {81,76,75}. Karena mempunyai *prefix* yang sama dengan lintasan TID 4 yaitu item 81 maka lintasan yang terbentuk pada TID 6 bisa dipadatkan pada lintasan TID 4. Setelah itu tambahkan *support count* pada item 81 menjadi dua karena telah dilewati sebanyak dua kali dan membentuk lintasan baru untuk item 76 dan 75 dengan *support count* bernilai satu.



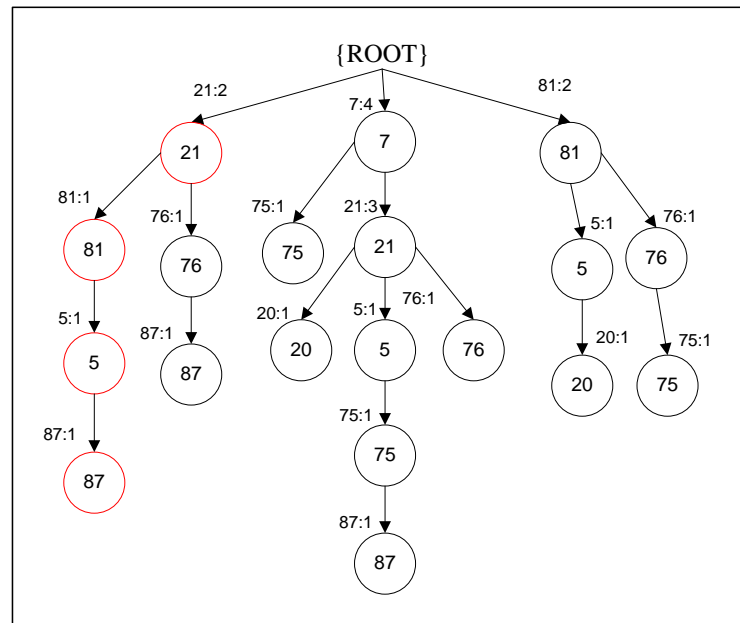
Gambar 0.6 Hasil pembentukan TID 6

Setelah pembacaan TID 6, selanjutnya adalah membaca TID 7 yaitu {21,76,87}. Karena prefix lintasannya tidak sama dengan lintasan sebelumnya, maka TID 7 dibuat lintasan baru.



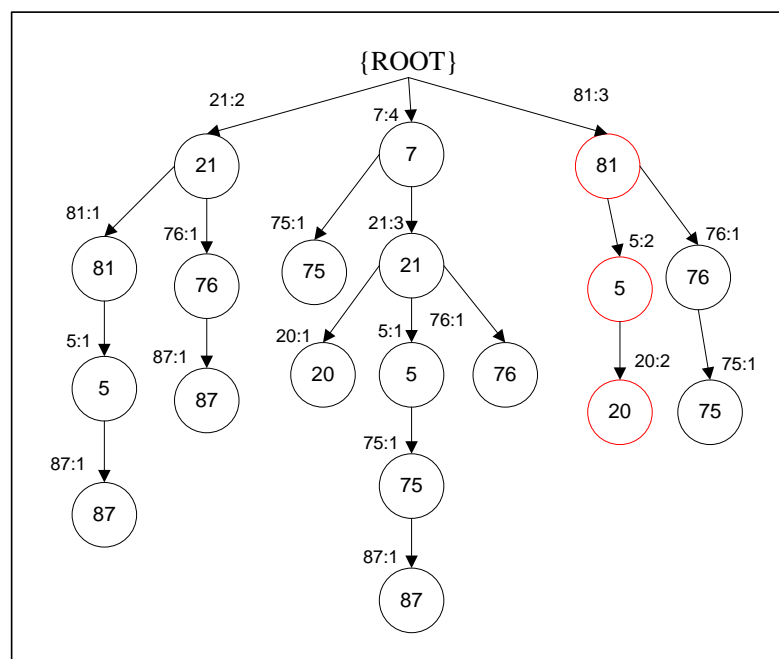
Gambar 0.7 Hasil pembentukan TID 7

Setelah pembacaan TID 7, selanjutnya adalah membaca TID 8 tetapi karena TID 8 item-item nya tidak memenuhi minimum *support*, maka TID 8 bernilai 0. Selanjutnya membaca TID 9 tetapi karena TID 9 item-item nya tidak memenuhi minimum *support*, maka TID 9 juga bernilai 0. Selanjutnya membaca TID 10 yaitu {21,81,5,87}. Karena mempunyai *prefix* yang sama dengan lintasan TID 6 yaitu item 21 maka lintasan yang terbentuk pada TID 10 bisa dipadatkan pada lintasan TID 6. Setelah itu tambahkan *support count* pada item 21 menjadi dua karena telah dilewati sebanyak dua kali dan membentuk lintasan baru untuk item 81, 5 dan 87 dengan *support count* bernilai satu.



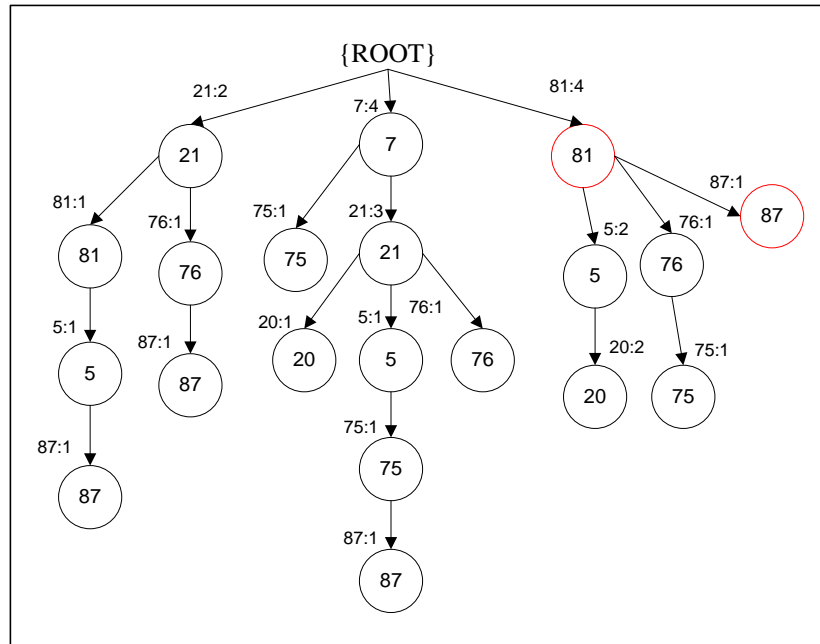
Gambar 0.8 Hasil pembentukan TID 10

Setelah pembacaan TID 10, selanjutnya adalah membaca TID 11 yaitu {81,5,20} tetapi karena TID 11 mempunyai *prefix* yang sama dengan TID 4 maka lintasan yang terbentuk sama dengan TID 4. Setelah itu tambahkan *support count* pada item 81 menjadi tiga karena telah dilewati sebanyak tiga kali, *support count* item 5 dan item 20 menjadi dua kali karena telah dilewati sebanyak dua kali.



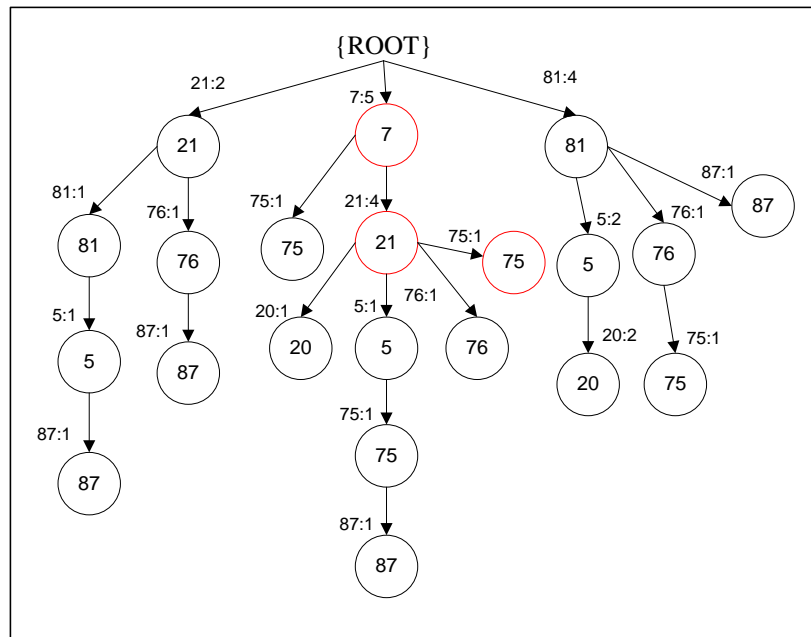
Gambar 0.9 Hasil pembentukan TID 11

Setelah pembacaan TID 11, selanjutnya adalah membaca TID 12 yaitu {81,87}. Karena TID 12 mempunyai *prefix* yang sama dengan TID 4 maka lintasan yang terbentuk sama dengan TID 4. Setelah itu tambahkan *support count* pada item 81 menjadi empat karena telah dilewati sebanyak empat kali, dan membentuk lintasan baru untuk item 87 dengan *support count* bernilai satu.



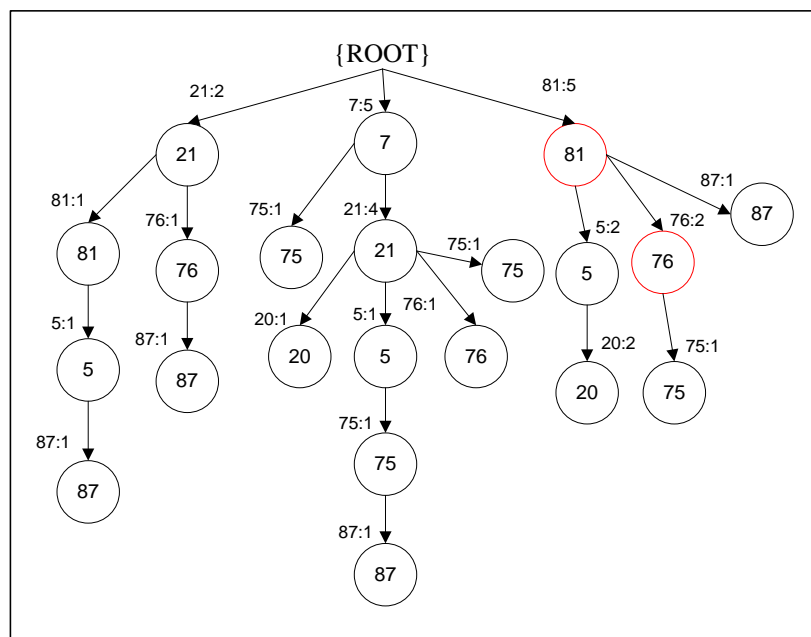
Gambar 0.10 Hasil pembentukan TID 12

Setelah pembacaan TID 12, selanjutnya adalah membaca TID 13 tetapi karena TID 13 item-item nya tidak memenuhi minimum *support*, maka TID 13 bernilai 0. Selanjutnya pembacaan TID 14 yaitu {7,21,75} tetapi karena TID 14 mempunyai *prefix* yang sama dengan TID 2 maka lintasan yang terbentuk sama dengan TID 2. Setelah itu tambahkan *support count* pada item 7 menjadi empat karena telah dilewati sebanyak empat kali dan *support count* pada item 21 menjadi dua karena telah dilewati sebanyak dua kali dan membentuk lintasan baru untuk item 75 dengan *support count* bernilai satu.



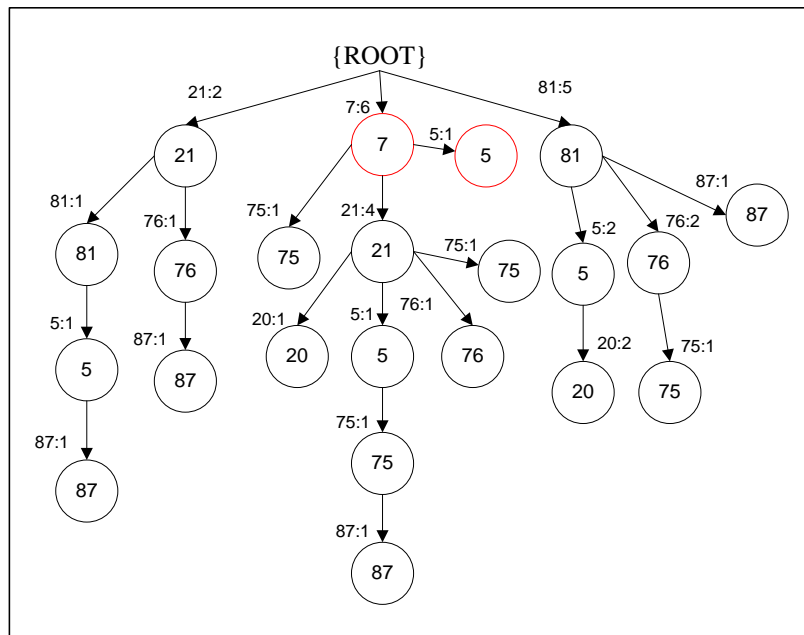
Gambar 0.11 Hasil pembentukan TID 14

Setelah pembacaan TID 14, selanjutnya adalah membaca TID 15 yaitu {81,76}. Karena TID 15 mempunyai *prefix* yang sama dengan TID 6 maka lintasan yang terbentuk sama dengan TID 6. Setelah itu tambahkan *support count* pada item 81 menjadi lima karena telah dilewati sebanyak lima kali dan *support count* pada item 76 menjadi dua karena telah dilewati sebanyak dua kali.



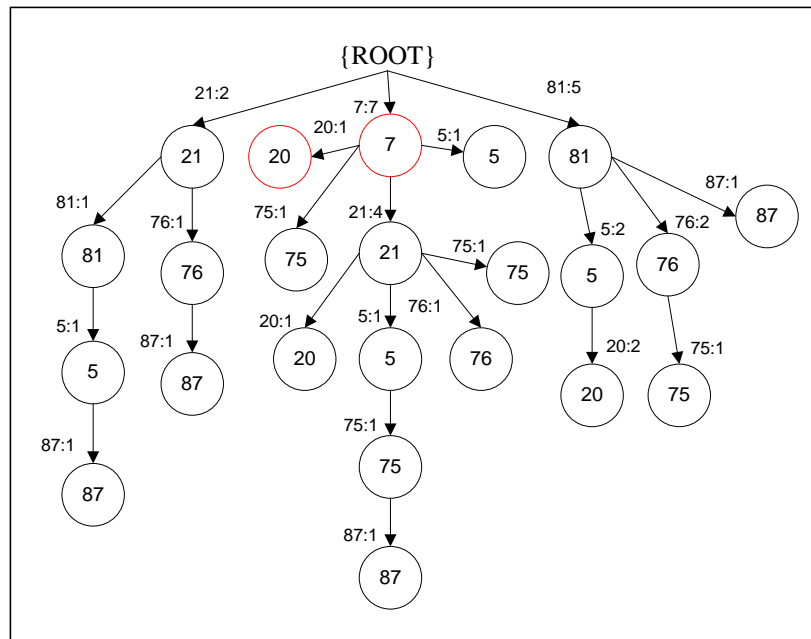
Gambar 0.12 Hasil pembentukan TID 15

Setelah pembacaan TID 15, selanjutnya adalah membaca TID 16 tetapi karena TID 16 item-item nya tidak memenuhi minimum *support*, maka TID 16 bernilai 0. Selanjutnya membaca TID 17 yaitu {7,5} tetapi karena TID 17 mempunyai *prefix* yang sama dengan TID 1 maka lintasan yang terbentuk sama dengan TID 1. Setelah itu tambahkan *support count* pada item 7 menjadi enam karena telah dilewati sebanyak enam kali dan membentuk lintasan baru untuk item 5 dengan *support count* bernilai satu.



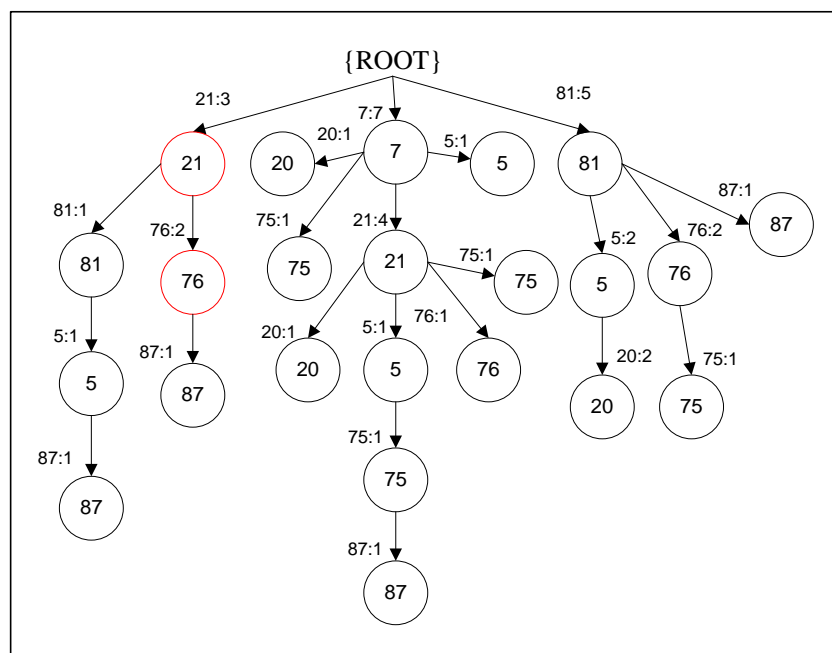
Gambar 0.13 Hasil pembentukan TID 17

Setelah pembacaan TID 17, selanjutnya adalah membaca TID 18 yaitu {7,20} tetapi karena TID 18 mempunyai *prefix* yang sama dengan TID 1 maka lintasan yang terbentuk sama dengan TID 1. Setelah itu tambahkan *support count* pada item 7 menjadi tujuh karena telah dilewati sebanyak tujuh kali dan membentuk lintasan baru untuk item 20 dengan *support count* bernilai satu.



Gambar 0.14 Hasil pembentukan TID 18

Setelah pembacaan TID 18, selanjutnya adalah membaca TID 19 yaitu {21,76} tetapi karena TID 19 mempunyai *prefix* yang sama dengan TID 7 maka lintasan yang terbentuk sama dengan TID 7. Setelah itu tambahkan *support count* pada item 21 menjadi tiga karena telah dilewati sebanyak tiga kali dan *support count* pada item 76 menjadi dua karena telah dilewati sebanyak dua kali.



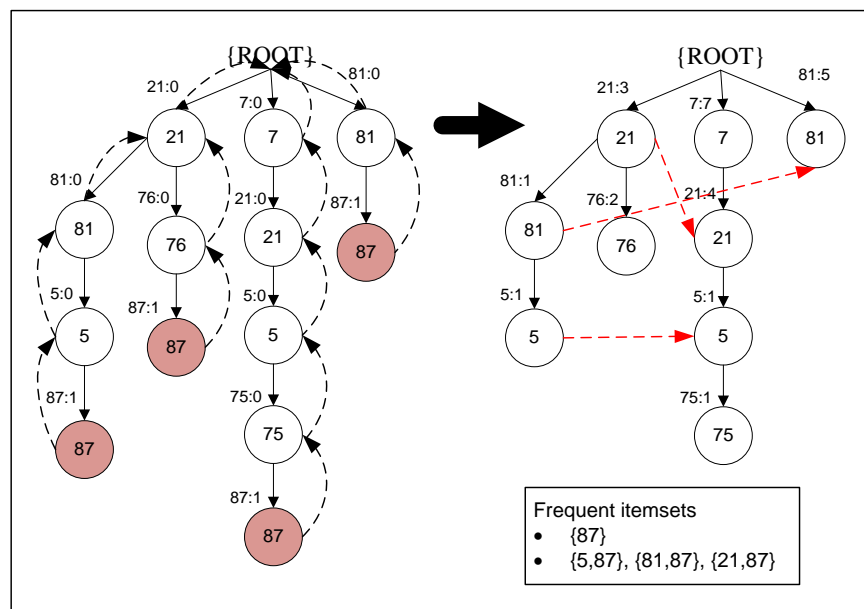
Gambar 0.15 Hasil pembentukan TID 19

Setelah pembacaan TID 19, selanjutnya adalah membaca TID 20 tetapi karena TID 19 item-item nya tidak memenuhi minimum *support*, maka TID 20 bernilai 0. Hasil pembentukan FP-Tree untuk contoh 20 data transaksi berakhir pada TID 19 seperti pada gambar 4.15.

Setelah pembuatan *FP-Tree* selesai, algoritma *FP-Growth* mencari semua *subsets* yang memungkinkan dengan cara membangkitkan *conditional FP-Tree* dan mencari *frequent itemset*.

1. Kondisi FP-Tree untuk item 87

Langkah awal yaitu mengekstrak semua lintasan yang berakhiran 87. Selain *path* 87 diberi nilai 0, hal ini dilakukan agar dapat mengetahui informasi berapa kali *item* yang lain dibeli bersamaan dengan *item* 87 dan bisa mengetahui *frequent itemsets* mana yang memenuhi syarat minimum *support*. Untuk lebih jelasnya, dapat dilihat pada gambar 4.16 dibawah ini.

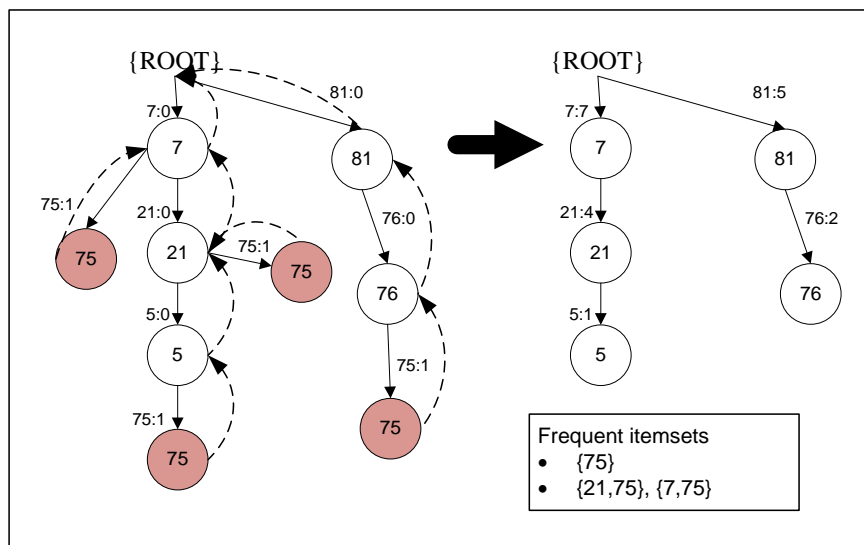


Gambar 0.16 Hasil pembentukan FP-Tree item 87

Setelah melakukan ekstrak untuk lintasan yang berakhiran 87, kemudian naikan path satu persatu sampai menuju ke *root* dan nilai path 87 di masukan pada setiap path yang dilintasi. Pada kondisi ini, item 7, 75 dan 76 akan dihilangkan

karena nilai kemunculannya secara bersamaan dengan item 87 hanya 1 kali. Sedangkan item 5, 21, dan 81 memiliki nilai kemunculan sebanyak lebih dari satu kali dan akan di simpan. Berdasarkan item yang memiliki kemunculan bersamaan dengan item 87 lebih dari satu kali tersebut maka subsets yang terbentuk adalah $\{87\}$, $\{5,87\}$, $\{81,87\}$, $\{21,87\}$.

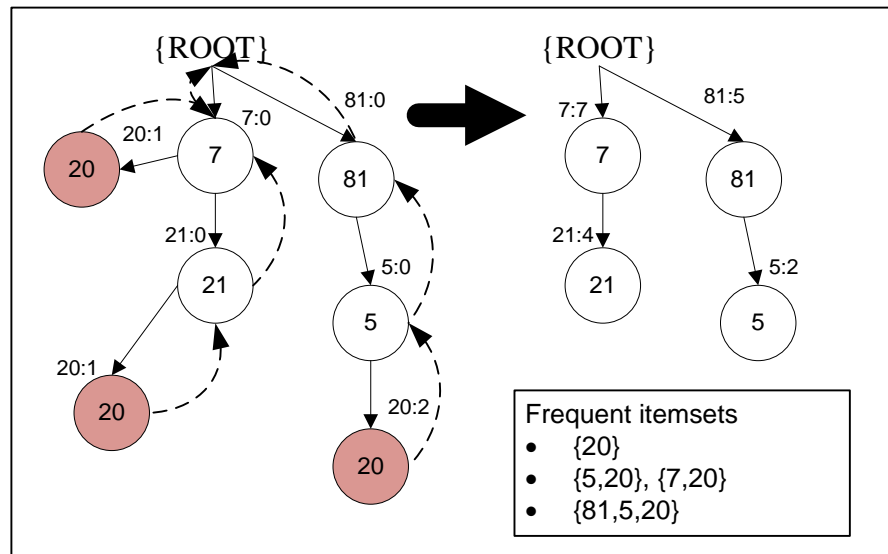
2. Kondisi FP-Tree untuk item 75



Gambar 0.17 Hasil pembentukan FP-Tree item 75

Setelah melakukan ekstrak untuk lintasan yang berakhiran 75 dan selain path 75 diberi nilai 0, kemudian naikan path satu persatu sampai menuju ke *root* dan nilai path 75 di masukan pada setiap path yang dilintasi. Pada kondisi ini, item 5, 76 dan 81 akan dihilangkan karena nilai kemunculannya secara bersamaan dengan item 75 hanya 1 kali. Sedangkan item 7 dan 21 memiliki nilai kemunculan sebanyak lebih dari satu kali dan akan di simpan. Berdasarkan item yang memiliki kemunculan bersamaan dengan item 75 lebih dari satu kali tersebut maka subsets yang terbentuk adalah $\{75\}$, $\{21,75\}$, $\{7,75\}$.

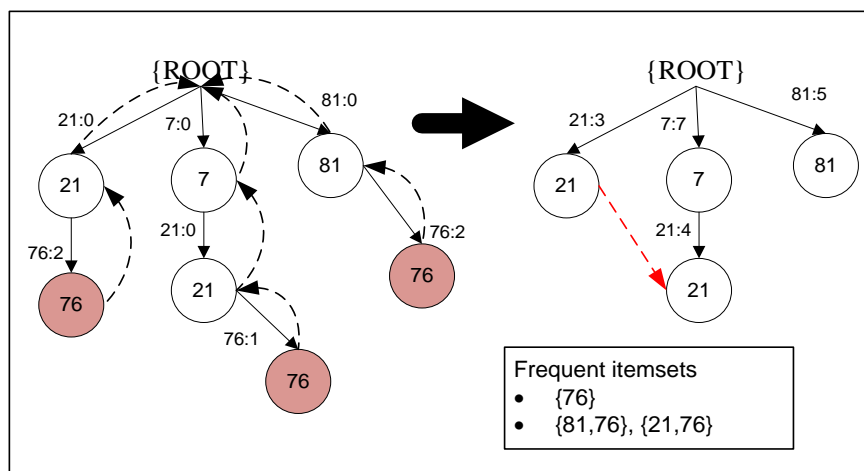
3. Kondisi FP-Tree untuk item 20



Gambar 0.18 Hasil pembentukan FP-Tree item 20

Setelah melakukan ekstrak untuk lintasan yang berakhiran 20 dan selain path 20 diberi nilai 0, kemudian naikan path satu persatu sampai menuju ke *root* dan nilai path 20 di masukan pada setiap path yang dilintasi. Pada kondisi ini item yang memiliki kemunculan bersamaan dengan item 20 lebih dari satu kali tersebut maka subsets yang terbentuk adalah {20}, {5,20}, {7,20}, {81,5,20}.

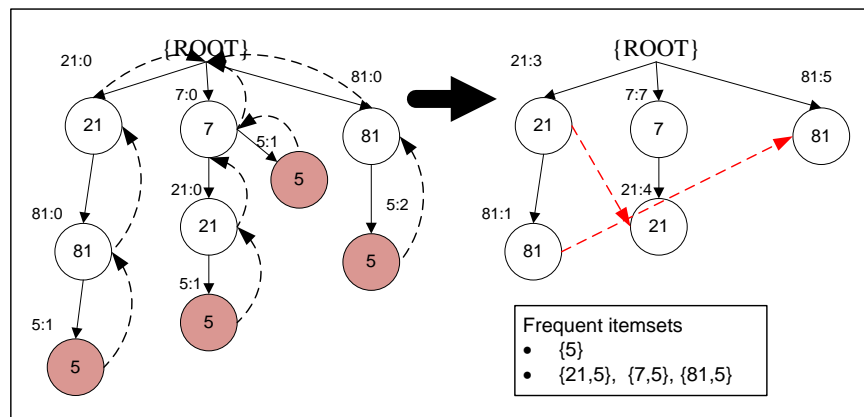
4. Kondisi FP-Tree untuk item 76



Gambar 0.19 Hasil pembentukan FP-Tree item 76

Setelah melakukan ekstrak untuk lintasan yang berakhiran 76 dan selain path 76 diberi nilai 0, kemudian naikan path satu persatu sampai menuju ke *root* dan nilai path 76 di masukan pada setiap path yang dilintasi. Pada kondisi ini, item 7 akan dihilangkan karena nilai kemunculannya secara bersamaan dengan item 76 hanya 1 kali. Sedangkan item 21 dan 81 memiliki nilai kemunculan sebanyak lebih dari satu kali dan akan di simpan. Berdasarkan item yang memiliki kemunculan bersamaan dengan item 76 lebih dari satu kali tersebut maka subsets yang terbentuk adalah {76}, {81, 76}, {21,76}.

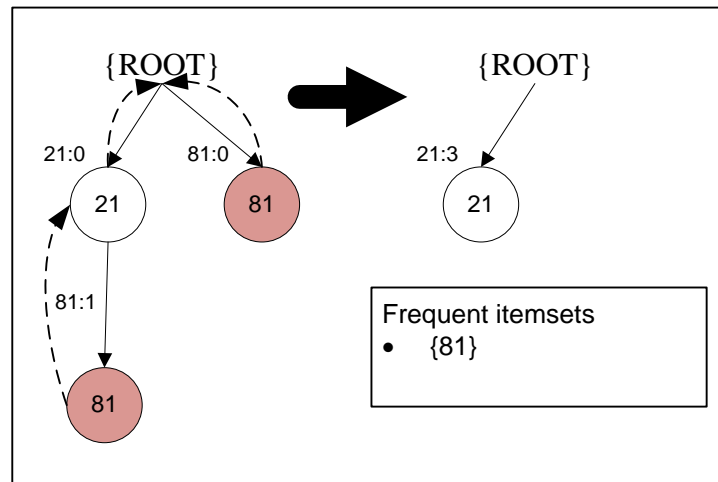
5. Kondisi FP-Tree untuk item 5



Gambar 0.20 Hasil pembentukan FP-Tree item 5

Setelah melakukan ekstrak untuk lintasan yang berakhiran 5 dan selain path 5 diberi nilai 0, kemudian naikan path satu persatu sampai menuju ke *root* dan nilai path 5 di masukan pada setiap path yang dilintasi. Pada kondisi ini, item 7, 21 dan 81 memiliki nilai kemunculan sebanyak lebih dari satu kali dan akan di simpan. Berdasarkan item yang memiliki kemunculan bersamaan dengan item 5 lebih dari satu kali tersebut maka subsets yang terbentuk adalah {5}, {21,5}, {7,5}, {81,5}.

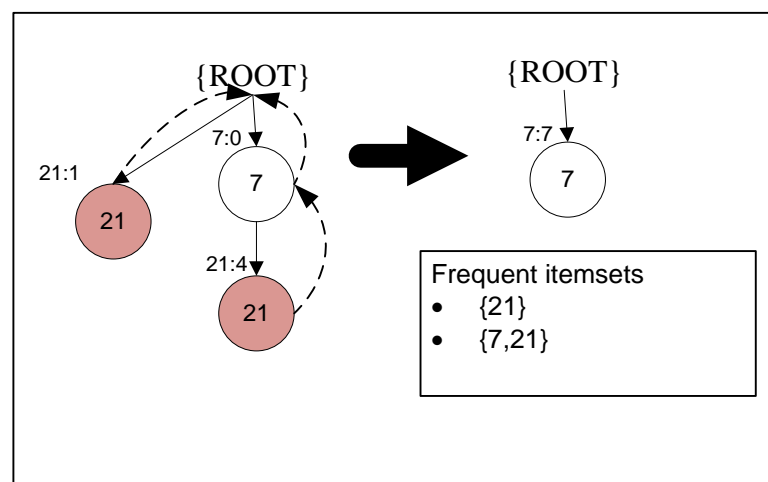
6. Kondisi FP-Tree untuk item 81



Gambar 0.21 Hasil pembentukan FP-Tree item 81

Setelah melakukan ekstrak untuk lintasan yang berakhiran 81 dan selain path 81 diberi nilai 0, kemudian naikkan path 81 satu persatu sampai menuju ke root dan nilai *path* 81 dimasukkan pada setiap path yang dilintasi. Pada kondisi ini, item 21 akan di buang karena kemunculannya bersama dengan item 81 hanya satu kali. Maka subsets yang terbentuk adalah {81}.

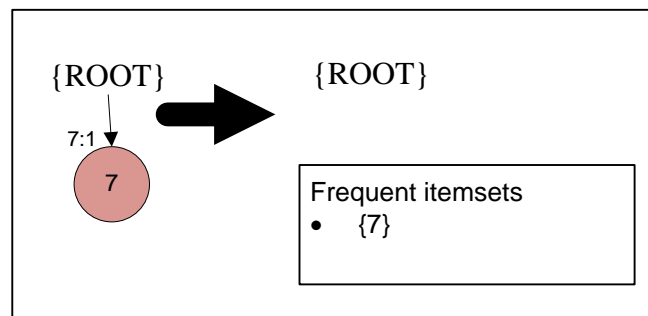
7. Kondisi FP-Tree untuk item 21



Gambar 0.22 Hasil pembentukan FP-Tree item 21

Setelah melakukan ekstrak untuk lintasan yang berakhiran 21 dan selain path 21 diberi nilai 0, kemudian naikan path 21 satu persatu sampai menuju ke root dan nilai *path* 21 dimasukkan pada setiap path yang dilintasi. Pada kondisi ini, subsets yang terbentuk adalah {21}, {7,21}.

8. Kondisi FP-Tree untuk item 7



Gambar 0.23 Hasil pembentukan FP-Tree item 7

Pada kondisi ini, lintasan yang berakhiran item 7 merupakan lintasan tunggal yang berdiri sendiri dan memiliki nilai frekuensi 7 sehingga frequent itemsets yang terbentuk hanya {7}.

Setelah memeriksa kondisi semua FP-Tree, maka didapatkan 21 frequent itemsets yang hasilnya dirangkum dalam tabel berikut:

Tabel 4.5 Hasil Frequent Itemset

Item Suffix	Frequent Itemsets
87	{87}, {5,87}, {81,87}, {21,87}
75	{75}, {21,75}, {7,75}
20	{20}, {7,20}, {5,20}, {81, 5, 20}
76	{76}, {81,76}, {21,76}
5	{5}, {21,5}, {7,5}, {81,5}
81	{81}
21	{21}, {7,21}
7	{7}

Berdasarkan 21 Frequent Itemset yang telah terbentuk diatas, tidak semua akan dihitung dalam proses selanjutnya, karena dalam menghasilkan *Association Rule* syarat *Frequent Itemsets* yang dihitung minimal memiliki 2 item dimana jika membeli A maka akan membeli B. Maka yang layak dihitung tingkat *confidence*-nya adalah 21 *subsets*, diantaranya :

Tabel 4.6 Tabel subset yang layak

Item Suffix	Frequent Itemsets
87	{5,87}, {81,87}, {21,87}
75	{21,75}, {7,75}
20	{7,20}, {5,20}, {81, 5, 20}
76	{81,76}, {21,76}
5	{21,5}, {7,5}, {81,5}
21	{7,21}

Setelah mendapatkan *subsets* yang layak dan memenuhi syarat, kemudian dihitung nilai *support* kemunculannya. Jika nilai *support frequent itemset* $\geq 15\%$ maka akan dihitung nilai *minimum confidence* yang telah ditentukan pada awal proses data mining. Semua kemungkinan pola yang dapat terbentuk dari masing-masing *subsets* akan dihitung nilai *confidence*-nya dan yang akan diambil adalah pola yang memiliki nilai *confidence* ≥ 0.60 . *Frequent itemsets* yang memenuhi nilai support yang telah ditentukan yaitu {21,87}, {7,75}, {21,76}, {81,5}, dan {7,21} untuk di cari kombinasinya dan dihitung nilai *confidence*-nya.

1. {21,87} terdapat 2 pola
 - $21 \rightarrow 87 = 3/7 = 0,428$
 - $87 \rightarrow 21 = 3/4 = 0,75$
2. {7,75} terdapat 2 pola
 - $7 \rightarrow 75 = 3/7 = 0,428$
 - $75 \rightarrow 7 = 3/4 = 0,75$

3. {21,76} terdapat 2 pola
 - $21 \rightarrow 76 = 3/7 = 0,428$
 - $76 \rightarrow 21 = 3/5 = 0,6$
4. {81,5} terdapat 2 pola
 - $81 \rightarrow 5 = 3/6 = 0,5$
 - $5 \rightarrow 81 = 3/5 = 0,6$
5. {81,5} terdapat 2 pola
 - $81 \rightarrow 5 = 3/6 = 0,5$
 - $5 \rightarrow 81 = 3/5 = 0,6$
6. {7,21} terdapat 2 pola
 - $7 \rightarrow 21 = 4/7 = 0,571$
 - $21 \rightarrow 7 = 4/7 = 0,571$

Dari perhitungan *confidence* terhadap pola yang terbentuk diatas, maka *Assosiattion Rule* yang memenuhi syarat $confidence \geq 0,60$ adalah :

1. $87 \rightarrow 21 = 0,75$ (jika konsumen membeli *bakery & pastry*, maka membeli *cookies* kering dengan nilai *confidence* 0,75)
2. $75 \rightarrow 7 = 0,75$ (jika konsumen membeli buah segar lokal, maka membeli *soft drink* dengan nilai *confidence* 0,75)
3. $76 \rightarrow 21 = 0,6$ (jika konsumen membeli buah segar import, maka membeli *cookies* kering dengan nilai *confidence* 0,6).
4. $5 \rightarrow 81 = 0,6$ (jika konsumen membeli *juice* buah, maka membeli susu segar & yoghurt dengan nilai *confidence* 0,6).

4.6 Implementasi Java

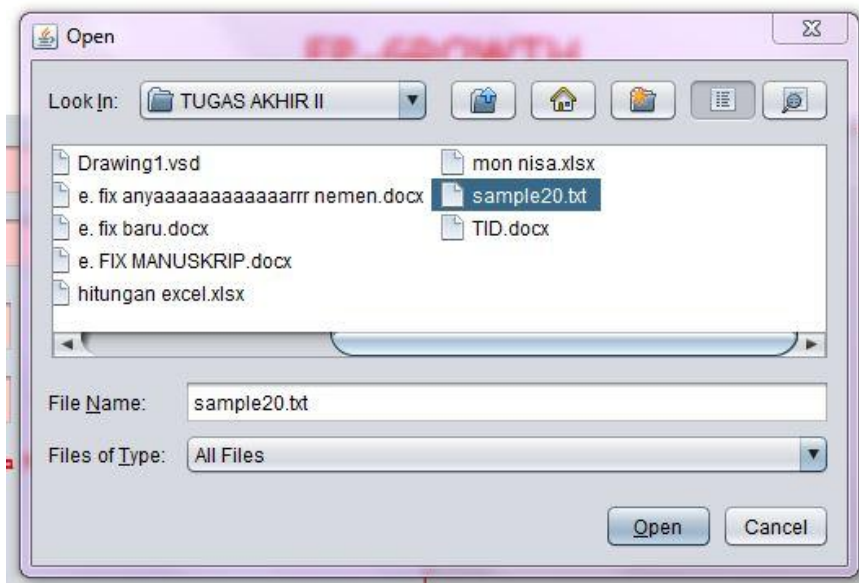
Pada tahapan implementasi, dataset transaksi penjualan yang telah melewati tahapan *preprocessing* dan *cleaning* akan diimplementasikan menggunakan java.

Implementasi *Market Basket Analysis* dengan menerapkan algoritma *FP-Growth* untuk menemukan pola pembelian *item* oleh konsumen, diaplikasikan kedalam bahasa pemrograman *java* menggunakan IDE Netbeans 7. Adapun tampilan dari aplikasi *Market Basket Analysis* seperti berikut



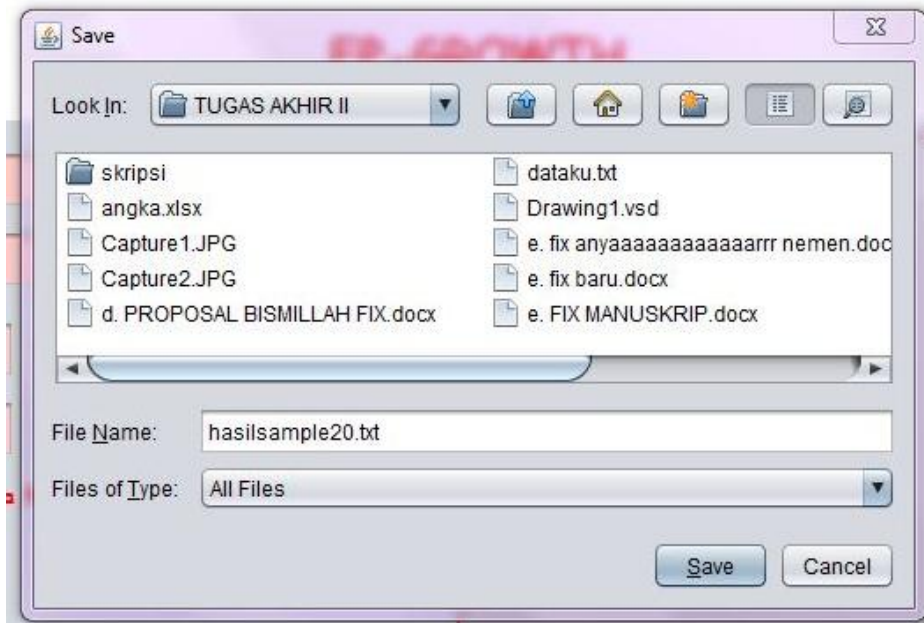
Gambar 0.24 Tampilan menu utama

Langkah selanjutnya, input data yang sesuai dengan pola untuk diproses dengan cara menekan buton browse pada input file. Pilih dan masukan file dalam bentuk .txt.



Gambar 0.25 Tampilan input file

Tekan button browse pada output file. Tentukan lokasi direktori penyimpanan hasil *generate rule*, ketik nama file diakhiri dengan .txt, contoh: hasilsample20.txt.



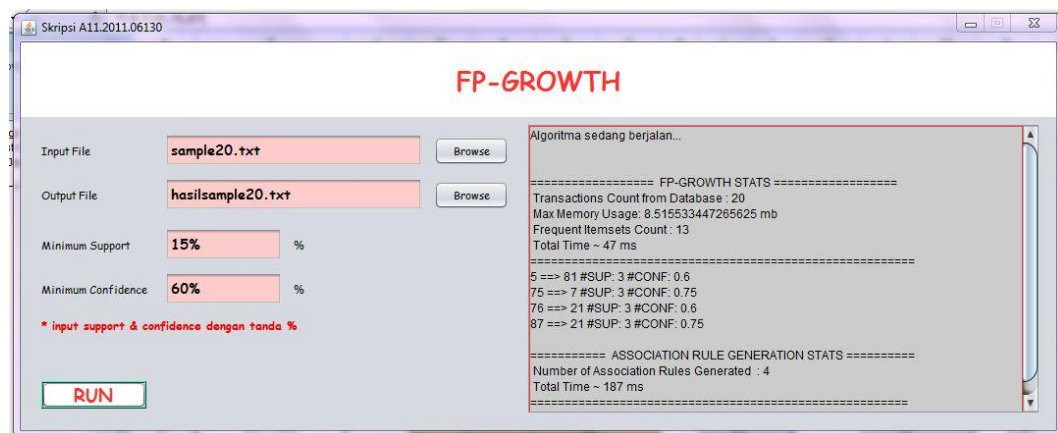
Gambar 0.26 Tampilan output file

Setelah data dan direktori penyimpanan hasil generate rule ditentukan, langkah selanjutnya adalah menginputkan nilai *support* dan *confidence*. Nilai *support* adalah frekuensi kemunculan pola *item* dalam seluruh transaksi yang ditentukan antara 0.0 (nilai persentase terkecil dari jumlah kemunculan pola *item*) sampai 1.0 (nilai persentase terbesar dari jumlah kemunculan pola *item*). Nilai *confidence* adalah nilai yang menghitung nilai kepercayaan dari pola *item* yang sudah memenuhi syarat nilai *support* yang ditentukan antara 0.0 (nilai persentase terkecil dari kepercayaan sebuah pola *item*) sampai 1.0 (nilai persentase terbesar dari kepercayaan sebuah pola *item*). Pada kasus ini, penulis menginput nilai minimum *support* dan *confidence* sama dengan nilai yang dipakai dalam perhitungan manual di atas.

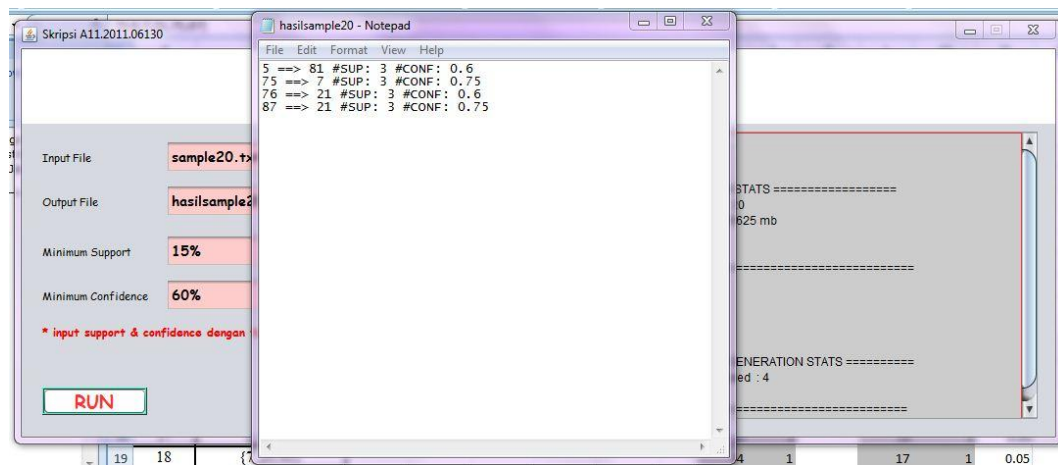


Gambar 0.27 Tampilan input min support dan min confidence

Setelah nilai min support dan min confidence telah diinputkan, langkah selanjutnya yaitu menjalankan program dengan cara menekan tombol run. Hasil yang muncul saat program di jalankan adalah seperti gambar 4.28

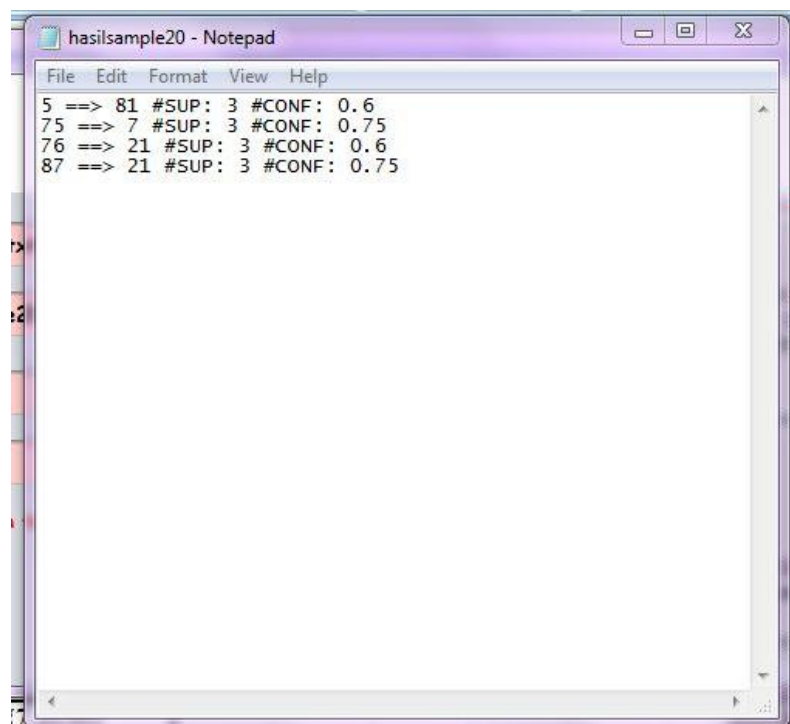


Gambar 0.28 Tampilan saat running program



Gambar 0.29 Tampilan hasil pada notepad

Pada tahap ini, diperlukan file data transaksi yang sudah disiapkan dengan menekan tombol *input file*, pada kasus ini penulis menginput data transaksi yang sama dengan contoh data yang telah dihitung secara manual menggunakan algoritma *FP-Growth* di atas. Bentuk data transaksinya bisa dilihat pada table 4.1 yang disimpan dalam bentuk .txt.



Gambar 0.30 Hasil aturan asosiasi yang terbentuk

Pada hasil di atas, dapat dibuktikan bahwa pola yang dihasilkan dari perhitungan menggunakan aplikasi, memiliki hasil pola yang sama dengan perhitungan manual.

4.7 Interpretation dan Evaluation

Hasil keseluruhan dari proses *generate rules* menggunakan aplikasi *Market Basket Analysis* terhadap data transaksi selama 1 bulan terakhir yaitu bulan Oktober 2014 pada Swalayan Gelael Candi sebanyak 1061 *record* dengan memberikan batasan nilai minimum *support* = 0.05 (5%) dan *confidence* = 0.3 (30%) .

Hasil di atas menunjukkan bahwa pola jika membeli salties maka membeli soft drink dengan nilai *support* = 10.27% dan nilai *confidence* = 45,6% dari 1061 data transaksi yang merupakan pola dengan nilai *support* dan *confidence* tertinggi.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil analisa proses data mining menggunakan algoritma FP-Growth didapatkan sebuah pola transaksi pembelian dengan frekuensi tertinggi sebagai berikut :

- a. Jika membeli salties maka membeli softdrink
- b. Jika membeli permen & jelly maka membeli softdrink
- c. Jika membeli wafer & biskuit maka membeli salties
- d. Jika membeli buah segar lokal maka buah segar import

Dari hasil analisa pola frekuensi yang dihasilkan oleh proses data mining pola transaksi yang dihasilkan dapat dijadikan sebuah rekomendasi dalam menentukan keputusan memberikan paket pembelian atau bundling dan bisa juga digunakan untuk tata letak layout atau promosi produk untuk meningkatkan penjualan di Swalayan Gelael Candi Semarang.

5.2. Saran

Dari hasil analisa data mining, penulis ingin memberikan saran yang sekiranya dapat dipertimbangkan untuk pengembangan data mining ini lebih lanjut :

- a. Data yang digunakan untuk diuji coba disarankan ditambah serta lebih memadatkan kriteria item agar dapat menghasilkan pola frekuensi tinggi dengan nilai keakuratan yang lebih tepat.
- b. Penelitian ini dapat dikembangkan dengan menggunakan metode aturan association rule data mining lainnya untuk dilakukan perbandingan.

DAFTAR PUSTAKA

- [1] Heru Dewantara, Purnomo Budi Santosa, and Nasir Widha Setyanto, "Perancangan Aplikasi Data Mining Dengan Algoritma Apriori Untuk Frekuensi Analisa Keranjang Belanja Pada Data Transaksi Penjualan," pp. 415-426, 2013.
- [2] Tri Lestari, "Analisis Keranjang Belanja Pada Data Transaksi Penjualan," 2009.
- [3] dkk Raorane A.A, "Association Rule - Extracting Knowledge Using Market Basket Analysis," *Recent Sciences*, 2012.
- [4] Yogi Yusuf, F. Rian Pratikto, and Gerry T, "Penerapan data Mining Dalam Penentuan Aturan Asosiasi Antar Jenis Item," *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, pp. E53-E56, Juni 2006.
- [5] Fatimah Fatihatul, Atje Setiawan, and Rudi Rosadi, "Asosiasi Data Mining Menggunakan Algoritma FP-Growth Untuk Market Basket Analysis," 2011.
- [6] Goldie Gunadi and Dana Indra Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-Growth)," *TELEMATIKA MKOM*, vol. IV, no. I, pp. 118-132, Maret 2012.
- [7] David Samuel, "Penerapan Stuktur FP-Tree Dan Algoritma FP-Growth Dalam Optimasi Penentuan Frequent Itemset," 2007.
- [8] Ririanti, "Implementasi Algoritma FP-Growth Pada Aplikasi Prediksi Persediaan Sepeda Motor," *Pelita Informatika Budi Darma*, vol. I, no. I, pp. 139-144, Maret 2014.
- [9] Kennedy Tampubolon, Hoga Saragih, and Bobby Reza, "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan," *INIT(Informasi dan Teknologi Ilmiah)*, vol. I, no. I, pp. 93-106, Oktober 2013.
- [10] Wiwit Agus Triyanto, "Association Rule Mining Untuk Penentuan Rekomendasi Promosi Produk," *SIMETRIS*, vol. 5, no. 2, pp. 121-126, November 2014.

- [11] Kusrini and Emha Taufiq Luthfi, *ALGORITMA DATA MINING*. Yogyakarta: ANDI, 2009.
- [12] D Larose, *Discovering Knowledge in Data : An Inroduction to Data Mining*. Hoboken: Wiley Intersciene, John Wiley and Sons, Inc, 2005.
- [13] Budi Santoso, *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu, 2007.
- [14] Dian Wirdasari and Ahmad Calam, "Penerapan Data Mining Untuk Mengelola Data Penempatan Buku Di Perpustakaan SMK TI pab Lubuk Pakam Dengan Metode Association Rule," *SAINTIKOM*, vol. 10, no. 2, pp. 137-150, 2011.
- [15] Kusrini and Emha Taufiq Luthfi, *Algoritma Data Mining*, 1st ed., Theresia Ari Prabawati, Ed. Yogyakarta: ANDI, 2009.
- [16] Erwin, "Analisis Market Basket Dengan Algoritma Aproiori Dan FP-Growth," *JURNAL GENERIC*, pp. 26-30, 2012.
- [17] Jingbo Yuan and Ding Shunlin, "Research And Improvement On Association Rule Algorithm Base On FP-Growth," 2012.