



Analisa Data Mining Menggunakan Frequent Pattern Growth pada Data Transaksi Penjualan PT Mora Telematika Indonesia untuk Rekomendasi Strategi Pemasaran Produk Internet

Harpa Erasmus Simanjuntak, Windarto*

Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

E-mail : ¹overlote19@gmail.com, ²windarto@budiluhur.ac.id

Email Penulis Korespondensi: windarto@budiluhur.ac.id

Abstrak—Pemanfaatan data transaksi penjualan yang banyak tersimpan dapat memberikan pengetahuan yang berguna dalam membuat kebijakan dan strategi bisnis bagi PT Mora Telematika Indonesia. Untuk mewujudkan hal tersebut dapat diterapkan dengan Market Basket Analysis. Association Rule adalah salah satu teknik data mining yang merupakan prosedur dalam Market Basket Analysis untuk mencari pengetahuan berupa pola pembelian konsumen. Pola ini dapat menjadi masukan dalam membuat kebijakan dan strategi bisnis. Suatu pola ditentukan oleh dua parameter, yaitu support (nilai penunjang) dan confidence (nilai kepastian). Dalam penelitian ini Market Basket Analysis menggunakan algoritma Frequent Pattern Growth (FP-Growth) untuk menemukan pola dengan menerapkan struktur data Tree atau disebut dengan FP-Tree. Salah satu pola yang dihasilkan dari analisis terhadap data transaksi penjualan pada kurun Januari 2018 sampai April 2018 yang menghasilkan 7 aturan asosiasi dengan nilai lift ratio tertinggi yaitu jika ada pemasangan OxygenHome 25 - Super Double maka akan ada pemasangan OxygenHome 15 - Super Double dengan lift ratio 4.59%, nilai support 3.125%, dan nilai confidence 0.67%.

Kata Kunci: Data Mining, Association Rule, FP-Growth, Tree pattern, FP Tree.

Abstract—Utilizing a lot of stored sales transaction data can provide useful knowledge in making policy and business strategy for PT Mora Telematika Indonesia. To realize the things can be applied with the Market Basket Analysis. Association Rule is a data mining technique which is a procedure in the Market Basket Analysis to find the knowledge of consumer purchase patterns. This pattern can be an input in making business policies and strategies. A pattern is determined by two parameters, which are support (supporting value) and confidence (value of certainty). In this study, the Market Basket Analysis used a Frequent Pattern Growth (FP-Growth) algorithm to find patterns by implementing TREE data structures or called FP-Tree. One of the patterns resulting from the analysis of data on sales transactions in the period of January 2018 to April 2018 is 7 Association rules with the highest lift ratio value is if there is an installation of OxygenHome 25-Super Double Then there will be installation OxygenHome 15-Super Double with elevator ratio 4.59%, support value of 3,125%, and confidence value 0.67%.

Keywords: Data Mining, Association Rule, FP-Growth, Tree Pattern, FP Tree.

1. PENDAHULUAN

Sebagai salah satu penyedia akses jaringan terbesar di Indonesia dengan salah satu produknya yaitu Oxygen, menuntut manajer penjualan PT Moratelindo harus dapat menentukan pola strategi pemasaran produk yang lebih efisien agar selalu unggul dalam bidang penjualan sehingga dapat memuaskan konsumen. Untuk mengatasi masalah dalam menemukan pola strategi pemasaran yang lebih efisien dapat dilakukan dengan menganalisa pola data hasil penjualan produk Oxygen.

Dalam melakukan penelitian ini, penulis merujuk pada beberapa literature terutama penelitian-penelitian yang terkait dengan pencarian pola asosiatif dengan menggunakan FP-Growth diantaranya sebagai berikut:

Banyaknya data permintaan pelanggan yang ditampung melalui sistem Customer Service Online System (CSOS) dalam melakukan pelayanan rutin kepada pelanggan di Kawasan Industri Batamindo (KIB). Data tersebut dapat diolah lebih lanjut menjadi suatu informasi yang berguna bagi Manajemen dalam pengambilan keputusan. Dengan adanya data Mining diharapkan dapat membantu PT. KIB dalam menggali informasi yang terkandung didalam suatu permintaan pelayanan menjadi sebuah pengetahuan (knowledge) yang baru. Association Rule, yaitu suatu prosedur dalam Market Analysis untuk mencari pengetahuan berupa hubungan kebiasaan pelanggan dalam meminta pelayanan berupa perbaikan gedung, informasi data dan konsultasi. Algoritma FP-Growth merupakan algoritma yang sangat efisien dalam mencari frequent itemset dalam sebuah kumpulan data dengan membangkitkan struktur prefix-tree. Data yang diolah adalah data permintaan pelanggan dari tahun 2016 dengan jumlah data 264 permintaan [1].

Masyarakat yang berkecimpung pada bidang bisnis sering membutuhkan suatu teknologi baru sebagai alat bantu mereka dalam menyelesaikan suatu pekerjaan. Misalnya suatu teknologi yang membantu mereka untuk memberikan analisis yang berhubungan dengan bisnis mereka. Salah satu bisnis yang membutuhkan teknologi seperti ini adalah bisnis retail. Bisnis retail membutuhkan teknologi tersebut misalnya untuk memberikan analisis apakah produk yang dijualnya sudah memenuhi kebutuhan pelanggan atau belum, produk apa saja yang biasa dibeli pelanggan, dan kebiasaan pelanggan membeli produk yang satu dengan yang lainnya. Jika terdapat produk yang tidak sesuai dengan kebutuhan pelanggan namun tetap dipertahankan bisa menyebabkan kerugian pada bisnis tersebut karena produk tidak laku dan menjadi kadaluwarsa. Sehingga perusahaan hanya menyediakan produk yang biasa dibeli oleh pelanggan untuk menghindari kerugian tersebut



dan mendapatkan keuntungan. Oleh karena itu, untuk memenuhi kebutuhan masyarakat tersebut penulis membangun aplikasi SMART–Sistem Informasi Manajemen Retail dengan menggunakan algoritma FP-Growth untuk menentukan produk yang harus disediakan oleh perusahaan berdasarkan Market Basket Analysis. Dari hasil pengujian dan analisa yang telah dilakukan, maka dapat disimpulkan bahwa pengimplementasian perancangan Market Basket Analysis dengan menggunakan FP-Growth di aplikasi SMART sudah berhasil. Rules yang dihasilkan di setiap bulan dengan $\text{min supp} = 0.002$ dan $\text{min conf} = 0.5$, jumlahnya hampir sama dan ada beberapa produk yang muncul di beberapa bulan yang berbeda. Semakin besar min supp dan min conf yang dimasukkan, maka menghasilkan rule yang semakin sedikit. Rata-rata dengan nilai $\text{min supp} = 0.006$ dan $\text{min conf} = 0.6$ tidak menghasilkan rules [2].

Pada koperasi karyawan PT. Phapros Semarang, data transaksi penjualan barang di toko koperasi yang terkumpul dan tersimpan dalam basis data hanya dipergunakan untuk menghasilkan laporan penjualan dan laporan laba rugi perusahaan saja. Padahal data tersebut dapat dimanfaatkan secara maksimal dengan diolah lebih lanjut sehingga didapatkan suatu informasi baru untuk pengambilan keputusan dalam strategi bisnis selanjutnya. Untuk mengatasi masalah tersebut, dibutuhkan suatu metode atau teknik yang dapat merubah tumpukan data menjadi informasi dan pengetahuan bermanfaat untuk pengambilan keputusan dan menentukan strategi bisnis koperasi karyawan seperti menentukan Cross-selling produk. Data mining merupakan solusi yang dapat digunakan untuk proses ekstraksi informasi pencarian pola atau trend yang diinginkan dalam basis data yang besar. Pola-pola ini dapat memberikan suatu analisis data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti untuk mendukung keputusan. Dalam penelitian ini, algoritma yang digunakan adalah Frequent Pattern-Growth (FP-Growth) yaitu pengembangan dari metode Apriori yang merupakan salah satu alternatif untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data dengan membangkitkan struktur data Tree atau disebut dengan Frequent Pattern Tree (FP-Tree). Algoritma FP-Growth merupakan salah satu metode yang tepat untuk diterapkan pada data transaksi penjualan yang digunakan untuk membantu manager dalam menganalisis pola pembelian konsumen sebagai pendukung keputusan dalam menentukan strategi bisnis yaitu cross-selling untuk meningkatkan pendapatan Koperasi Karyawan PT. Phapros secara berturut-turut dengan menjual produk tambahan yang berasal dari produk utama yang dibeli oleh pelanggan. Strategi bisnis cross-selling dapat dilakukan dengan cara menyusun layout yang baik berdasarkan pada pola pembelian item dengan mendekati item yang memiliki nilai confidence tinggi serta memberikan paket diskon (product bundling) terhadap pola pembelian item yang memiliki nilai confidence tinggi namun memiliki nilai support yang kecil [3].

Dari tinjauan beberapa penelitian diatas dapat disimpulkan bahwa *state of the art* dari penelitian ini adalah untuk menemukan pola penjualan produk oxygen di PT Mora Telematika Indonesia berdasarkan data penjualan produk Oxygen. Sehingga hasil dari penelitian ini dapat menemukan pola penjualan produk oxygen yang dapat digunakan sebagai strategi pemasaran produk Oxygen yang lebih efisien agar penjualan produk tersebut selalu unggul sehingga dapat memuaskan konsumen. Hal tersebut dikarenakan saat ini masih ditemukan bahwa belum semua area menggunakan produk Oxygen. Maka untuk mengatasi masalah dalam menemukan pola strategi pemasaran yang lebih efisien dapat dilakukan dengan menganalisa pola atau hubungan asosiatif data hasil penjualan produk Oxygen.

Pencarian pola atau hubungan asosiatif dari data yang berskala besar sangat erat kaitannya dengan penambangan data (*data mining*). *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar[4]. Istilah lain yang memiliki makna sama dengan *data mining*, yaitu *Knowledge discovery in databases* (KDD), ekstraksi pengetahuan (*knowledge extraction*), Analisa data/pola (*data/pattern analysis*), kecerdasan bisnis (*business intelligence*), data *archaeology*, serta *data dredging* [5]. Pencarian pola atau hubungan asosiatif dari data yang berskala besar sangat erat kaitannya dengan *data mining*. Aturan asosiatif (*association rule*) adalah salah satu teknik utama dalam *data mining* dan merupakan bentuk paling umum yang dipakai dalam menemukan pola (*pattern*) dari suatu kumpulan data.

2. METODOLOGI PENELITIAN

Algoritma *FP-Growth* merupakan salah satu alternatif yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. Pada algoritma *FP-Growth* menggunakan konsep pembangunan *tree*, yang biasa disebut *FP-Tree* dalam pencarian *frequent itemsets*, bukan menggunakan *generate candidate* seperti yang dilakukan pada algoritma Apriori. Dengan menggunakan konsep tersebut, algoritma *FP-Growth* menjadi lebih cepat daripada algoritma Apriori [2].

2.1 Association Rule

Association rules merupakan salah satu metode yang bertujuan mencari pola yang sering muncul diantara banyak transaksi, dimana setiap transaksi terdiri dari beberapa *item*. Sehingga metode ini akan menemukan hubungan antar *item*. Ini mungkin memerlukan pembacaan data transaksi secara berulang-ulang dalam jumlah data transaksi yang besar untuk menemukan pola hubungan yang berbeda-beda. Metode dasar analisis asosiasi terbagi menjadi tiga tahap:



- a. Analisa pola frekuensi tinggi.

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. *Support* merupakan suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *itemset* dari keseluruhan transaksi. Nilai *support* 1 *item* diperoleh dari rumus (1), sedangkan nilai *support* 2 *item* diperoleh dari rumus (2).

$$Supp A = \frac{\text{jumlah transaksi A}}{\text{jumlah total transaksi}} \tag{1}$$

$$Supp (A, B) = \frac{\text{jumlah transaksi A dan B}}{\text{jumlah total transaksi}} \tag{2}$$

- b. Pembentukan aturan asosiatif

Setelah semua pola frekuensi tinggi ditemukan, kemudian dicari aturan asosiatif yang memenuhi *confidence* dengan menghitung *confidence* aturan *assosiatif* $A \rightarrow B$. *Confidence* merupakan suatu ukuran yang menunjukkan hubungan antara dua *item* secara kondisional (Contoh: seberapa sering *item* B dibeli jika pembeli membeli *item* A) yang dinotasikan dengan:

$$conf (A \rightarrow B) = \frac{\text{jumlah transaksi A dan B}}{\text{jumlah transaksi A}} \tag{3}$$

- c. Proses perhitungan *lift ratio*

Lift ratio adalah suatu ukuran (parameter) untuk mengetahui kekuatan aturan asosiasi (*association rule*) yang telah dibentuk dari nilai *support* dan *confidence*. Nilai *lift ratio* biasanya digunakan sebagai penentu apakah aturan asosiasi valid atau tidak valid.

$$expected\ conf = \frac{\text{jumlah transaksi B}}{\text{jumlah total transaksi}} \tag{4}$$

$$lift\ ratio = \frac{confidence}{expected\ confidence} \tag{5}$$

2.2 Frequent Pattern Growth (FP-Growth)

FP-Growth merupakan salah satu algoritma yang termasuk dalam *association rule mining*. Algoritma *FP-Growth* dibagi menjadi tiga langkah utama, yaitu:

- 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 atau sama dengan *minimum support count* akan dibangkitkan dengan *conditional FP-tree*.

- c. Tahap pencarian *frequent itemset*.

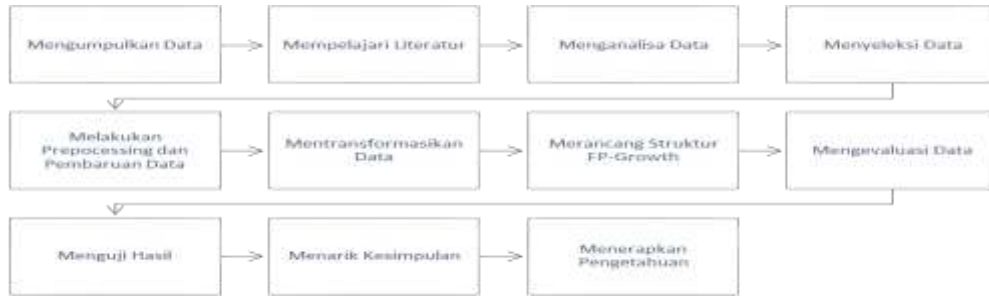
Apabila *Conditional FP-tree* merupakan lintasan tunggal (*single path*), maka didapatkan *frequent itemset* dengan melakukan kombinasi *item* untuk setiap *conditional FP-Tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan *FP-growth* secara rekursif.

2.3 Bahan Riset/Data

Tabel 1. Tabel Atribut Data Transaksi Penjualan Produk *Oxygen*

Kode Atribut	Keterangan
Transaksi Id	Id dari setiap transaksi sebagai kode unik untuk digunakan dalam algoritma <i>FP-Growth</i> untuk mencari <i>FP-Tree</i>
Id Item produk	Id dari setiap produk <i>Oxygen</i> yang dipakai costumer
Id Area	Adalah nama dari setiap produk <i>Oxygen</i> yang dipakai
Tanggal Transaksi	Id dari setiap area yang menggunakan produk <i>Oxygen</i>
	Merupakan tanggal aktivasi dari produk <i>Oxygen</i>

Tabel 1 diatas memperlihatkan data transaksi penjualan produk *Oxygen* yang akan dipakai untuk mendapatkan hasil dari pengujian aplikasi yang dibuat.

**2.4 Tahapan Penelitian****Gambar 1.** Kerangka Kerja Penelitian

Langkah-langkah penelitian dapat dijelaskan sebagai berikut:

1. Mengumpulkan Data

Teknik pengumpulan data yang digunakan adalah teknik studi dokumen, yaitu metode pengumpulan data yang tidak ditujukan langsung kepada subjek penelitian. Teknik studi dokumen yang digunakan peneliti adalah teknik studi dokumen sekunder, dimana dilakukan pengamatan terhadap objek data dengan tujuan mendapatkan korelasi antara persediaan barang dengan kebutuhan pelanggan.

2. Mempelajari Literatur

Metode pengumpulan data dengan studi kepustakaan dilakukan untuk memperoleh informasi dan menggali teori-teori tentang teknik penambangan data. Dalam penelitian ini penulis mempelajari berbagai literatur dari beberapa jurnal yang membahas tentang aplikasi penambangan data terutama menggunakan metode algoritma *FP-Growth* yang berguna sebagai referensi untuk sistem yang akan dibangun.

3. Menganalisa Data

Pada tahap ini penulis melakukan analisa data menggunakan tahapan KDD (*Knowledge Discovery in Database*). Di mana tahapan KDD ini meliputi: Seleksi Data, *Preprocessing* dan Pembersihan Data, Transformasi Data, *Data Mining*, Evaluasi/Interpretation.

4. Menyeleksi Data (*Data Selection*)

Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target, atau memfokuskan pada subset variabel (*sampel data*) dimana penemuan (*discovery*) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional *Preprocessing*.

5. Melakukan *Preprocessing* dan Pembersihan Data (*Cleaning Data*)

Pre processing dan *Cleaning Data* dilakukan dengan membuang data yang tidak konsisten dan *noise*, duplikasi data, memperbaiki kesalahan data dan bisa diperkaya dengan data eksternal yang relevan.

6. Mentransformasi Data (*Data Transformation*)

Proses ini mentransformasikan atau menggabungkan data kedalam yang lebih tepat untuk melakukan proses *mining* dengan cara melakukan peringkasan (*agregasi*).

7. Merancang Struktur *K Means Clustering*

Proses *Data Mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode atau algoritma tertentu sesuai dengan tujuan dari proses KDD secara keseluruhan. Proses KDD terdiri dari *data cleaning*, *data integration*, *data selection*, *data transformation*, *data mining*, *pattern evaluation* dan *knowledge presentation*. Pada tahap awal akan dilakukan *data cleaning* dan *data integration* pada data penjualan produk *Oxygen* sehingga data tersebut dapat digunakan untuk mencari pola asosiasinya. Proses awal ini dilakukan secara manual menggunakan Microsoft Excel. Lalu untuk proses selanjutnya seperti *data selection*, *data transformation* dan *data mining* akan dilakukan pada aplikasi yang akan dibuat. Sedangkan *pattern evaluation* dan *knowledge presentation* baru dapat dilakukan setelah aplikasi selesai dibangun karena proses ini membutuhkan hasil dari alat uji tersebut.

Algoritma yang digunakan dalam penelitian ini menggunakan *FP-Growth*. *FP-Growth* merupakan salah satu algoritma yang termasuk dalam *association rule mining*. Algoritma *FP-Growth* dibagi menjadi tiga langkah utama, yaitu:

1. Tahap pembangkitan *Conditional Pattern Base*.

Conditional Pattern Base merupakan *subdatabase* yang berisi lintasa *prefix (prefix path)* dan pola akhiran (*suffix pattern*). Pembangkitan *conditional 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 atau sama dengan *minimum support count* akan dibangkitkan dengan *conditional FP-Tree*.

3. Tahap pencarian *frequent itemset*.



Apabila hasil dari *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.

8. Mengevaluasi Data (*Evaluation Data*)

Tahap selanjutnya yaitu mengevaluasi data. Tahap ini digunakan untuk menguji kualitas dari data apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta sebelumnya.

9. Menguji Hasil

Tahap pengujian hasil adalah suatu teknik yang digunakan untuk menentukan bahwa penerapan metode yang digunakan telah mampu memecahkan masalah. Pada penelitian ini dilakukan pengujian dengan menggunakan aplikasi *data mining* yang dikembangkan oleh penulis. Berdasarkan pengetahuan yang diperoleh dari pola-pola yang terbentuk, maka pola yang terbentuk tersebut dapat dipresentasikan ke dalam bentuk visualisasi.

10. Menarik Kesimpulan

Setelah melakukan analisa data menggunakan tahapan KDD maka dapat ditarik suatu kesimpulan apakah informasi yang ditemukan berkesesuaian dengan fakta sebelumnya atau tidak, sehingga akan diperoleh pengetahuan baru.

11. Menerapkan Pengetahuan.

Langkah selanjutnya yaitu menerapkan pengetahuan yang diperoleh, sehingga diharapkan dapat membantu pimpinan organisasi/perusahaan dalam proses pembuatan keputusan yang akan datang.

3. HASIL DAN PEMBAHASAN

Sumber data utama yang digunakan dalam penelitian ini berasal dari data transaksi pada kurun 02 Januari 2018 sampai dengan 25 April 2018 dengan transaksi Id 1 sampai transaksi Id 96 sebanyak 490 data, yang didapatkan dari departemen IT di PT Mora Telematika Indonesia.

Tabel 2. Contoh Tabel Transaksi

CUSTOMER_NO	24100178
CUSTOMER_NAME	AHMAD SETIAWAN
CUSTOMER TIPE_NAME	Perumahan
AREA_NAME	CIRACAS
ADDRESS_HOMEPASS	Jl. PUCUNG 3 9B 012 004 CIRACAS KRAMAT JATI JAKARTA DKI JAKARTA
PACKAGE_NAME	OxygenHome 15
CUSTOMER_STATUS	Active
TANGGAL_TRANSAKSI	02/01/2018
PACKAGE PRICE	262900

Pada data transaksi penjualan terdapat banyak atribut. Proses pembersihan (*cleaning*) dilakukan untuk memilih atribut data yang menjadi fokus penelitian serta menghapus atribut yang tidak diperlukan, yakni atribut **PACKAGE_NAME**, **TANGGAL_TRANSAKSI** dan **AREA_NAME**. Setelah proses ini, selanjutnya disimpan untuk diproses ke tahap selanjutnya.

Tabel 3. Tabel Contoh Data Setelah Proses Transformasi Data

Transaksi Id	1
Id Produk	3
Nama Produk	OxygenHome 15
Id Area	8
Tgl Transaksi	02/01/2018

3.1 Penerapan Algoritma FP-Growth

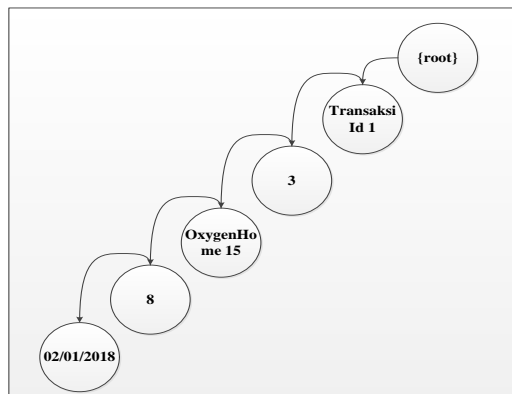
Pada tahap ini, *FP-Tree* dibuat untuk digunakan bersamaan dengan algoritma *FP-Growth* untuk menentukan *frequent itemsets*. Tabel 4 memperlihatkan contoh beberapa data yang diambil dari 6 transaksi.

Tabel 4. Contoh Data Transaksi Yang Sudah Ditransformasi

Transaksi Id	Id Produk	Nama Produk	Id Area	Tgl Transaksi
1	3	OxygenHome 15	8	02/01/2018
1	5	OxygenHome 15–Double Bandwidth	3	02/01/2018
2	5	OxygenHome 15–Double Bandwidth	32	03/01/2018
2	4	OxygenHome 25–Double Bandwidth	22	03/01/2018
3	5	OxygenHome 15–Double Bandwidth	15	04/01/2018

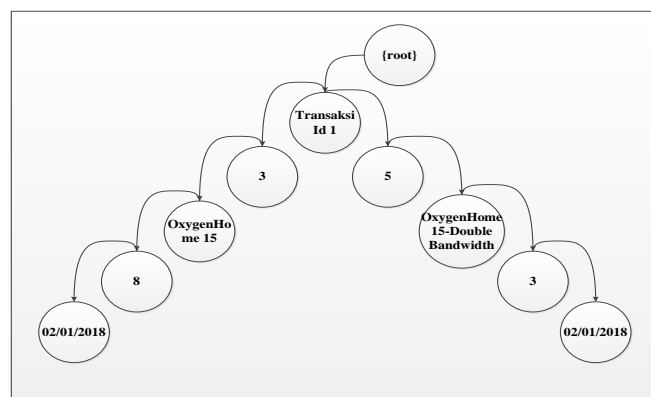


Transaksi Id	Id Produk	Nama Produk	Id Area	Tgl Transaksi
3	5	OxygenHome 15-Double Bandwidth	3	04/01/2018



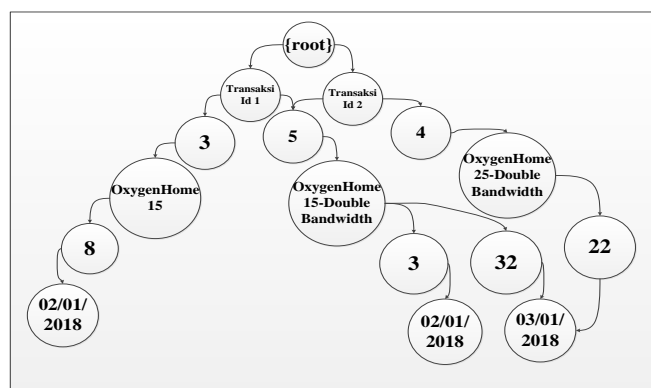
Gambar 2. Hasil pembentukan FP-Tree setelah pembacaan Transaksi Id 1

Dari gambar 2 dapat dijelaskan bahwa setelah *item* disusun ulang berdasarkan *F-list*, kemudian dilakukan penelusuran *database* yang kedua, yaitu membaca setiap transaksi diawali dengan membaca Transaksi Id 1 untuk membuat *FP-Tree*. Transaksi Id 1 {3, OxygenHome15, 8, 02/01/2018} akan membuat simpul sehingga terbentuk lintasan {} → Transaksi Id 1 → 3 → OxygenHome15 → 8 → 02/01/2018 dengan *support count* awal bernilai satu.



Gambar 3. Hasil pembentukan FP-Tree setelah pembacaan Id produk 5

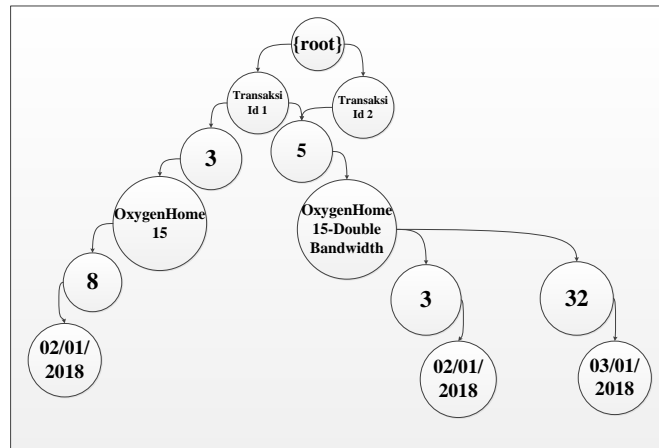
Dari gambar 3 dapat dijelaskan bahwa setelah pembacaan Transaksi Id 1 pada *item* pertama, maka selanjutnya membaca Transaksi Id 1 pada Id produk 5 yaitu {5, OxygenHome15-DoubleBandwidth, 3, 02/01/2018} sehingga membentuk lintasan {} → Transaksi Id 1 → 5 → OxygenHome15 → 3 → 02/01/2018 dengan *support count* awal bernilai satu juga. Walaupun Id produk 5 ada pada Transaksi Id 1, tetapi karena *prefix* transaksinya tidak sama, maka Id produk 5 ini tidak dipadatkan ke lintasan Id produk 3.



Gambar 4. Hasil pembentukan FP-Tree setelah pembacaan Transaksi Id 2

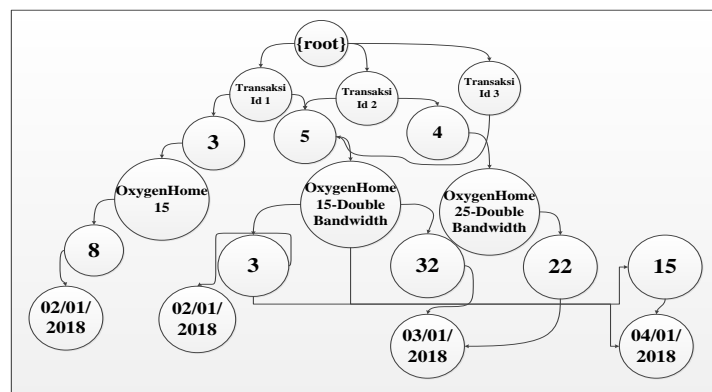


Dari gambar 4 dapat dijelaskan bahwa setelah pembacaan Transaksi Id 1, maka selanjutnya membaca Transaksi Id 2 yaitu {5, OxygenHome15-DoubleBandwidth, 32, 03/01/2018} sehingga membentuk lintasan{}→ 5 → OxygenHome15-DoubleBandwidth → 32 → 03/01/2018. Karena memiliki *prefix* yang sama dengan lintasan Transaksi Id 1 yaitu Id produk 5 dan Oxygen15-DoubleBandwidth, maka lintasan Transaksi Id 2 bisa dipadatkan pada lintasan Transaksi Id 1. Setelah itu tambahkan *support count* Id produk 5 dan OxygenHome15-DoubleBandwidth menjadi dua karena telah dilewati sebanyak dua kali, sedangkan Id Area dan Tanggal Transaksi masing-masing bernilai *support count* satu.



Gambar 5. Hasil pembentukan FP-Tree setelah pembacaan Transaksi Id 2 Id produk 4

Dari gambar 5 dapat dijelaskan bahwa setelah pembacaan Transaksi 2, maka selanjutnya membaca Transaksi Id 2 dengan Id produk 4 yaitu {4, OxygenHome15-DoubleBandwidth, 22, 03/01/2018} sehingga membentuk lintasan{}→ 4 → OxygenHome25-DoubleBandwidth → 22 → 03/01/2018 dengan *support count* awal bernilai satu juga. Walaupun Id produk 4 ada pada Transaksi Id 2, tapi karena *prefix* transaksinya tidak sama dengan lintasan Id produk 5, maka Id produk 4 ini tidak dipadatkan ke lintasan Id produk 5. Tetapi, pada tanggal transaksi bisa dipadatkan pada lintasan Tanggal Transaksi sebelumnya dengan menambahkan *support count* satu pada Tanggal Transaksi.



Gambar 6. Hasil pembentukan FP-Tree setelah pembacaan Transaksi Id 3

Dari gambar 6 dapat dijelaskan bahwa setelah pembacaan Transaksi Id 2, maka selanjutnya membaca Transaksi Id 3 yaitu {5, OxygenHome15-DoubleBandwidth, 15, 04/01/2018} sehingga membentuk lintasan{}→ 5 → OxygenHome15-DoubleBandwidth → 15 → 04/01/2018. Karena memiliki *prefix* yang sama dengan lintasan Transaksi Id 1 yaitu Id produk 5 dan Oxygen15-DoubleBandwidth, maka lintasan Transaksi Id 3 bisa dipadatkan pada lintasan Transaksi Id 1. Setelah itu tambahkan *support count* id produk 5 dan OxygenHome15-DoubleBandwidth menjadi tiga karena telah dilewati sebanyak tiga kali, sedangkan Id Area dan Tanggal masing-masing bernilai *support count* satu. Untuk gambar lintasan ini, disatukan dengan gambar Transaksi Id 3 produk selanjutnya karena kedua produk Transaksi Id 3 lintasannya dipadatkan pada lintasan sebelumnya.

Masih pada gambar 6, setelah pembacaan Transaksi Id 3, maka selanjutnya membaca Transaksi Id 3 produk selanjutnya {5, OxygenHome15-DoubleBandwidth, 3, 04/01/2018} sehingga membentuk lintasan{}→ 5 → OxygenHome15-DoubleBandwidth → 3 → 04/01/2018. Karena memiliki *prefix* yang sama dengan lintasan Transaksi Id 2 dan Transaksi 3 yaitu Id produk 5 dan Oxygen15-DoubleBandwidth dan Id area 3, maka lintasan Transaksi Id 3 bisa dipadatkan pada lintasan Transaksi Id 2 dan Transaksi Id 3. Setelah itu tambahkan *support count* id produk 5 dan OxygenHome15-DoubleBandwidth menjadi empat karena telah dilewati sebanyak empat



kali, sedangkan Id Area menjadi dua karena dilewati sebanyak dua kali dan Tanggal Transaksi menjadi dua karna dilewati sebanyak dua kali.

3.2 Pengujian

Metode untuk pengujian aplikasi adalah metode pengujian *dataset* dan uji validasi. Untuk pengujian *dataset*, pengujian dilakukan dengan mencoba mengkombinasikan nilai *minimum support* dan *minimum confidence* untuk melihat nilai *lift ratio* yang dihasilkan. Hasil dari kombinasi tersebut juga dapat dilihat nilai *count support* dan *benchmark* nya. Untuk pengujian validasi, hasil dari aplikasi yang dibuat dengan menerapkan algoritma *FP-Growth* akan dibandingkan dengan hasil dari aplikasi pada penelitian sebelumnya dengan menerapkan algoritma apriori.

3.2.1 Pengujian Beta

Pengujian beta dilakukan dengan mencoba mengkombinasikan nilai *minimum support* dan *minimum confidence* untuk melihat nilai *lift ratio* yang dihasilkan dan spesifikasi rekomendasi dengan perantara *user* dari PT Mora Telematika Indonesia tepatnya di Departemen IT. Hasil dari kombinasi tersebut juga dapat dilihat nilai *count support* dan *benchmark*.

Min Support Absolut: 30 %
Min Confidence: 50 %

HEADER / F-LIST

NAMA ITEM	N(A)
OxygenHome 15	95
OxygenHome 15 - Double Bandwidth	267
OxygenHome 25 - Double Bandwidth	45
OxygenHome 50 - Double Bandwidth	10
OxygenHome 100 - Double Bandwidth	11
OxygenHome 25 - Double Bandwidth	95
OxygenHome 25	15
OxygenHome 100	2
OxygenHome 25 - Double Bandwidth	267
OxygenHome 15 - Super Double	14
OxygenITC 8 Mbps	14
OxygenHome 25 - Super Double	14
OxygenHome 15	267

Gambar 7. Aturan Asosiasi Dengan Minimum Support 0.3 dan Confidence 0.5

Pada tahap pengujian ini dilakukan dengan menggunakan nilai *minimum support* dan nilai *minimum confidence* sesuai batasan yaitu nilai *minimum support* 30% (0.3) dan nilai *minimum confidence* 50% (0.5) menggunakan beberapa data transaksi sebagai *dataset* dengan batasan yang telah ditentukan. Hasil dari perhitungan ini pada tahap ini menghasilkan aturan asosiasi yang dapat dilihat pada gambar 8.

Support dan Confidence

RULE	COUNT	SUPPORT	CONFIDENCE
Rule 1 & 4 => 0	25	10.4166666666667	2.50
Rule 1 & 4 => 2	2	10.4166666666667	
Rule 1 => 2	0	14.5833333333333	
Rule 0 => 1	2	3.125	0.67
Rule 0 => 1	16	278.125	
Rule 3 => 1	12	0	inf
Rule 4 => 1	10	46.875	

Benchmark dan Lift Ratio

ITEM	COUNT	SUPPORT	CONFIDENCE	Frekuensi Item Consequent	BENCHMARK	LIFTRATIO
OxygenHome 25 - Doble Bandwidth & OxygenHome 15 - Doble Bandwidth => OxygenHome 100 - Doble Bandwidth	25	10.4166666666667	2.50	267	2.78125	0.90
OxygenHome 15 - Super Double & OxygenHome 25 - Doble Bandwidth => OxygenHome 15 - Doble Bandwidth	2	10.4166666666667		11	0.114583333333333	1.75
OxygenHome 15 - Super Double & OxygenHome 100 - Doble Bandwidth	0	14.5833333333333		11	0.114583333333333	3.75
OxygenHome 25 - Super Double & OxygenHome 15 - Super Double	2	3.125	0.67	14	0.145833333333333	4.50
OxygenHome 15 - Doble Bandwidth & OxygenHome 15 - Super Double	16	278.125		14	0.145833333333333	0.41
OxygenHome 15 & OxygenHome 15 - Super Double	12	0	inf	14	0.145833333333333	0.00
OxygenHome 25 - Doble Bandwidth & OxygenHome 15 - Super Double	10	46.875		14	0.145833333333333	1.61

Gambar 8. Spesifikasi Aturan Asosiasi Yang Dihasilkan



Setelah didapatkan aturan asosiasi yang dapat dilihat pada gambar 7, tahapan berikutnya adalah melakukan perhitungan *count rule*, *support rule* dan *confidence rule* serta *benchmark* dan *lift ratio*.

Dari gambar 8 dapat dilihat bahwa hasil dari perhitungan ini didapatkan berdasarkan pengujian beta terhadap data transaksi yang telah dilakukan sehingga dapat disimpulkan bahwa sistem yang dibangun telah memberikan beberapa hasil berupa spesifikasi aturan asosiasi yaitu sebagai berikut:

- Jika ada pemasangan OxygenHome 25 - Doble Bandwidth dan OxygenHome 15 - Doble Bandwidth maka akan ada pemasangan OxygenHome 100 - Doble Bandwidth dengan *lift ratio* 0.90 %
- Jika ada pemasangan OxygenHome 15 - Super Double dan OxygenHome 25 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Doble Bandwidth dengan *lift ratio* 1.75%
- Jika ada pemasangan OxygenHome 15 - Super Double maka akan ada pemasangan OxygenHome 100 - Doble Bandwidth dengan *lift ratio* 3.75%
- Jika ada pemasangan OxygenHome 25 - Super Double maka akan ada pemasangan OxygenHome 15 - Super Double dengan *lift ratio* 4.59%
- Jika ada pemasangan OxygenHome 15 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Super Double dengan *lift ratio* 0.41%
- Jika ada pemasangan OxygenHome 15 maka akan ada pemasangan OxygenHome 15 - Super Double dengan *lift ratio* 0.00
- Jika ada pemasangan OxygenHome 25 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Super Double dengan *lift ratio* 1.51 %

3.2.2 Pengujian Hasil

No	Item 1	Item 2	Item 3	Jumlah	Support
Itemset 3 yang lolos:					
<div style="background-color: #e0ffe0; padding: 5px;"> ✓ Success Proses mining selesai </div>					
Confidence dari itemset 3					
No	X → Y	Support X U Y	Support X	Confidence	Keterangan
Confidence dari itemset 2					
No	X → Y	Support X U Y	Support X	Confidence	Keterangan
1	OxygenHome 15 - Doble Bandwidth => OxygenHome 25 - Doble Bandwidth	30.43	82.61	36.64	Tidak Lolos
2	OxygenHome 25 - Doble Bandwidth => OxygenHome 15 - Doble Bandwidth	30.43	39.43	100.00	Lolos
3	OxygenHome 15 - Doble Bandwidth => OxygenHome 15	30.43	82.61	36.64	Tidak Lolos
4	OxygenHome 15 => OxygenHome 15 - Doble Bandwidth	30.43	87.83	63.64	Lolos
Rule Asosiasi yang terbentuk:					
No	X → Y	Confidence	Nilai Uji IR	Kategori rule	
1	OxygenHome 25 - Doble Bandwidth => OxygenHome 15 - Doble Bandwidth	100.00	1.21	kategori positif	
2	OxygenHome 15 => OxygenHome 15 - Doble Bandwidth	83.64	0.77	kategori negatif	

Gambar 9. Pengujian Hasil Algoritma Apriori

Pengujian hasil dilakukan untuk mengetahui aplikasi yang telah dibangun dapat memberikan hasil sesuai dengan yang diharapkan. Pada penelitian sebelumnya dengan menggunakan algoritma apriori, telah dilakukan uji coba pada data produk yang sama, dimana hasilnya dapat dilihat pada **gambar 9**.

Benchmark dan Lift Ratio

ITEM	COUNT	SUPPORT	CONFIDENCE	Frekuensi Item Consequent	BENCHMARK	LIFTRATIO
OxygenHome 25 - Doble Bandwidth & OxygenHome 15 - Doble Bandwidth => OxygenHome 100 - Doble Bandwidth	25	10.4166666666667	2.50	267	2.78125	0.90
OxygenHome 15 - Super Double & OxygenHome 25 - Doble Bandwidth => OxygenHome 15 - Doble Bandwidth	2	10.4166666666667		11	0.114583333333333	1.75
OxygenHome 15 - Super Double & OxygenHome 100 - Doble Bandwidth	6	14.5833333333333		11	0.114583333333333	3.75
OxygenHome 25 - Super Double & OxygenHome 15 - Super Double	2	3.125	0.67	14	0.145833333333333	4.59
OxygenHome 15 - Doble Bandwidth & OxygenHome 15 - Super Double	16	270.125		14	0.145833333333333	0.41
OxygenHome 15 & OxygenHome 15 - Super Double	12	0	inf	14	0.145833333333333	0.00
OxygenHome 25 - Doble Bandwidth & OxygenHome 15 - Super Double	10	46.875		14	0.145833333333333	1.51

Hasil

SPEKIFIKASI	LIFTRATIO
Jika ada pemasangan OxygenHome 25 - Doble Bandwidth dan OxygenHome 15 - Doble Bandwidth maka akan ada pemasangan OxygenHome 100 - Doble Bandwidth	0.90
Jika ada pemasangan OxygenHome 15 - Super Double dan OxygenHome 25 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Doble Bandwidth	1.75
Jika ada pemasangan OxygenHome 15 - Super Double maka akan ada pemasangan OxygenHome 100 - Doble Bandwidth	3.75
Jika ada pemasangan OxygenHome 25 - Super Double maka akan ada pemasangan OxygenHome 15 - Super Double	4.59
Jika ada pemasangan OxygenHome 15 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Super Double	0.41
Jika ada pemasangan OxygenHome 15 maka akan ada pemasangan OxygenHome 15 - Super Double	0.00
Jika ada pemasangan OxygenHome 25 - Doble Bandwidth maka akan ada pemasangan OxygenHome 15 - Super Double	1.51

Gambar 10. Pengujian Hasil Algoritma FP-Growth

