

IMPLEMENTASI DATA MINING PADA PENJUALAN PRODUK DI CV CAHAYA SETYA MENGGUNAKAN ALGORITMA FP-GROWTH

Wahyu Nur Setyo¹, Sukma Wardhana²

Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana^{1,2}

Jl. Raya Meruya Selatan, Kembangan, Jakarta, 11650

E-mail : wahyunursetyo17@gmail.com, sukma@mercubuana.ac.id

ABSTRACT

At this time the growth of data occurs rapidly and fast long with the use of computer systems in various transactions. But this increasingly large volume of data has no meaning if it is not processed into a knowledge which can be done by data mining. Association rule or what is known as market based analysis is one type of data mining implementation. This study aims to find patterns of transaction data in the CV Cahaya Setya retail industry by using a Frequent Pattern Growth algorithm also known as FP-Growth algorithm. FP-Growth aims to find all the set items that can be retrieved (often found) from the transaction database as efficiently as possible. The results of this study show that the pattern on the database of consumer transactions at CV Cahaya Setya retail industry the FP-Growth algorithm and implement it in the application

Keywords: Data mining, retail industry, association rule, FP-growth algorithm

ABSTRAK

Pada saat ini pertumbuhan data terjadi dengan cepat dan pesat seiring dengan penggunaan sistem komputer dalam berbagai transaksi. Tetapi volume data yang semakin besar ini tidak memiliki arti jika tidak diolah menjadi sebuah pengetahuan dimana hal ini dilakukan oleh data mining. Association rule atau yang dikenal sebagai market based analysis merupakan salah satu tipe implementasi dari data mining. Beberapa penelitian telah dilakukan untuk mencari pola pada data transaksi dengan tipe implementasi association rule. Penelitian ini bertujuan untuk mencari pola dari data transaksi di industri retail yaitu pada CV Cahaya Setya dengan menggunakan algoritma Frequent Pattern Growth atau yang dikenal pula sebagai FP-Growth. FP-Growth bertujuan untuk menemukan semua itemset yang dapat diambil (yang sering ditemukan) dari basis data transaksi seefisien mungkin. Hasil penelitian ini memperlihatkan bahwa pola pada basis data transaksi konsumen pada industri retail yaitu pada CV Cahaya Setya dapat ditemukan dengan menggunakan algoritma FP-Growth kemudian mengimplementasikannya pada aplikasi penjualan produk di cv cahaya setya menggunakan fp growth

Kata Kunci: data mining, market based analysis, association rule (aturan sosiasi), frequent pattern growth (FP-Growth), industri retail

1. PENDAHULUAN

Pada saat ini industri perdagangan semakin berkembang dengan data transaksi yang meningkat jumlahnya setiap saat. Terhadap kondisi tersebut dibutuhkan sebuah metode atau teknik yang dapat mengubah hamparan data tersebut menjadi sebuah informasi atau pengetahuan (*knowledge*) yang bermanfaat untuk mendukung pengambilan keputusan dalam bisnis. Teknologi yang dapat digunakan untuk mewujudkannya adalah *data mining*. Proses pengolahan atas data dengan jumlah yang besar dengan menggunakan teknik dan metode yang beragam. Dari beberapa metode yang sering digunakan dalam teknologi *data mining* adalah metode asosiasi atau *association rule mining*. Di dalam bidang usaha distributor dan penjualan berbentuk *retail*, metode *association rule mining* ini lebih dikenal dengan istilah analisa keranjang belanja (*market basket analysis*). Analisa terhadap kecenderungan perilaku konsumen dalam melakukan belanja barang [1]. Sumber data dari *market basket analysis* antara lain dapat bersumber dari transaksi kartu kredit, kupon diskon, atau panggilan pesanan pelanggan. *Market basket analysis* umumnya dimanfaatkan sebagai titik awal pencarian pengetahuan dari suatu transaksi data ketika kita tidak mengetahui pola spesifik apa yang kita cari. Kebutuhan *market basket analysis* berawal dari keakuratan dan manfaat yang dihasilkannya dalam wujud aturan asosiasi (*association rules*), pola-pola dari data yang memiliki keterkaitan dalam basis data [2]. Sebagai perusahaan retail, CV Cahya Setya memiliki begitu banyak data transaksi, dengan tipe pembelian konsumen yang beragam. Membeli barang dengan jumlah satuan, ataupun membeli jenis barang secara berpasangan sesuai dengan kebutuhannya, contohnya membeli perangkat TV dan unit antenanya sekaligus ataupun membeli rak dengan atribut di dalamnya. Banyaknya data transaksi penjualan tersebut semakin meningkat dengan setiap harinya. Terdapat beberapa algoritma yang dapat digunakan untuk menyusun *Association rule*. Pada penelitian ini, penulis akan menggunakan algoritma *FP-Growth*, sebuah metode dalam data mining untuk mencari Frequent *itemset* Tanpa menggunakan *candidate generation*. Pembangunan data menggunakan struktur FP-Tree untuk mengolah *database transaksi* [3]. *FP-Growth* menggunakan strategi *divide-conquer* sehingga algoritma ini hanya membutuhkan dua kali *scanning database*. Tujuan yang ingin di capai dalam penelitian ini antara lain adalah: (1) membuat fitur informasi penjualan CV Cahya Setya dengan menerapkan metode *data mining association rule* menggunakan algoritma *FP-Growth*, (2) Membangun aplikasi berbasis desktop dengan menggunakan bahasa pemrograman PHP.

2. METODE

A. Data Mining

Kerangka proses *data mining* tersusun atas tiga tahapan, yaitu pengumpulan data (*data collection*), transformasi data (*data transformation*), dan analisis data (*data analysis*) [4].

B. Algoritma Fp-Growth

Frequent Pattern Growth (FP-Growth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data [5]

Tahapan yang akan dilakukan untuk mendapatkan *frequent itemset* menggunakan algoritma *FP-growth* terbagi menjadi tiga langkah, antara lain:

- Pembangkitan *Conditional Pattern Base*.
- Pembangkitan *Conditional FP- tree*.
- Pencarian frequent itemset.

C. FP-TREE

FP- TREE adalah struktur penyimpanan data yang dibangun dengan memetakan setiap data transaksi kedalam setiap lintasan tertentu. *FP-tree* digunakan dalam mencari pola frequent dengan batas ambang minimum *support count* ξ dengan menggunakan algoritma *FP-growth*. Jika *support*

dari pola tersebut tidak kurang dari konstanta ξ (batas ambang minimum support) yang telah ditetapkan, maka pola tersebut dapat dikatakan sebagai frequent pattern (sering muncul). Adapun definisi dari FP-tree adalah [6]:

- a. FP-tree terbentuk dari sebuah akar dengan label *null*;
- b. Setiap simpul mengandung informasi: label item (jenis item yang direpresentasikan), *support count* (jumlah lintasan transaksi, dan pointer (penghubung simpul dengan label sama).

D. HYPER TEXT MARKUP LANGUAGE (HTML)

Hyper Text Mark Up Language merupakan bahasa asli dari www, yang telah menjadi bahasa standard untuk menampilkan data di internet Perkembangan html sangat pesat, saat ini versi terakhir dari html telah mencapai html [7].

E. HYPERTEXT PROTOCOL (PHP)

Pengertian PHP menurut php adalah bahasa pemrograman web berupa script yang dapat di integrasikan dengan HTML. PHP merupakan bahasa scripting yang open source dan digunakan untuk membuat website yang dinamis dan powerfull[8].

Pemrograman yang berjalan pada server memang banyak sekali, Setiap Program memiliki kelebihan dan kekurangan.Saat ini banyak website yang menggunakan PHP sebagai dasar pengolahan data.Beberapa keunggulan yang dimiliki program PHP (MADCOMS 2004) di antaranya adalah:

- 1) PHP memiliki tingkat akses yang lebih cepat
- 2) PHP memiliki tingkat lifecycle yang cepat sehingga selalu mengikuti perkembangan teknologi internet
- 3) PHP memiliki tingkat keamanan yang tinggi
- 4) PHP mampu berjalan di beberapa server yang ada misalnya Apache, Microsoft IIS, PWS, AOLserver, phttpd, fhttpd, dan xitami
- 5) PHP mampu berjalan di LINUX sebagai platform sistem operasi utama bagi PHP
- 6) PHP mendukung ke beberapa database yang sudah ada
- 7) PHP bersifat free atau gratis

F. MYSQL

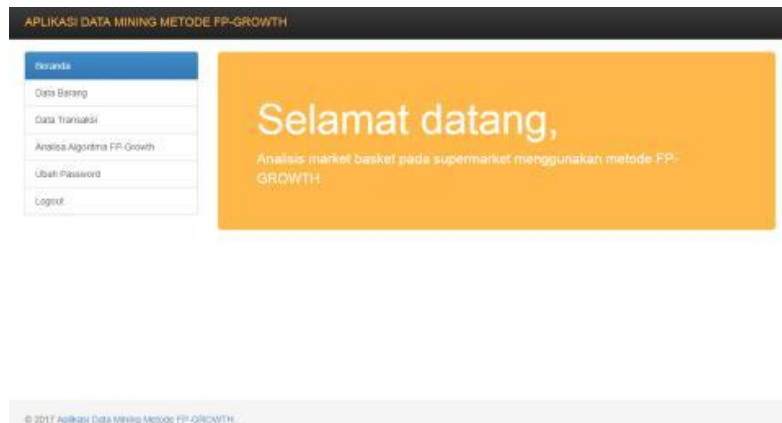
MySQL adalah sebuah software yang Open Source, sehingga bebas dipakai dan dimodifikasi oleh semua orang. Setiap orang dapat mendownload MySQL dari internet dan menggunakannya tanpa harus membayar [9]. MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL *Database Management System* atau DBMS[10].

3. HASIL DAN PEMBAHASAN

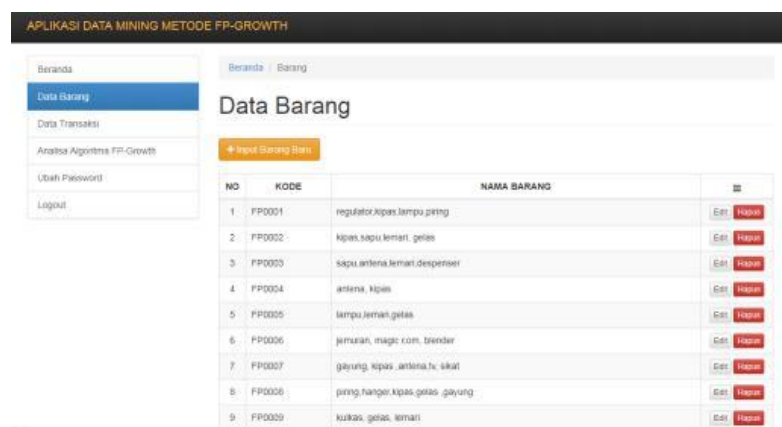
A. Hasil



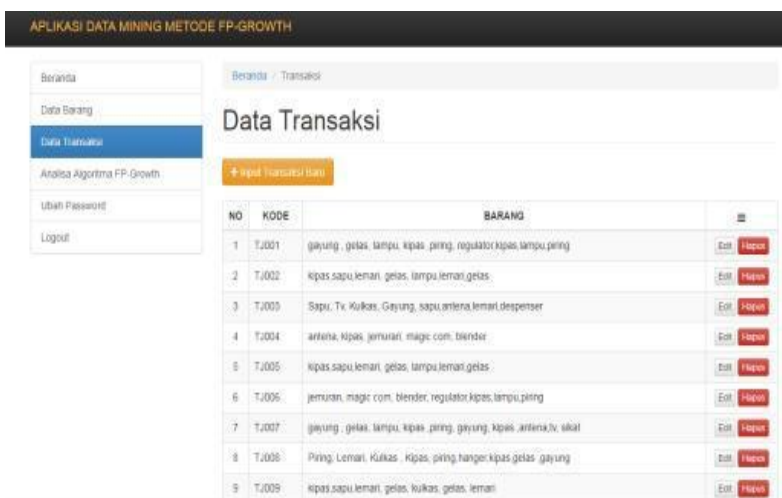
Gambar 1. Tampilan Home



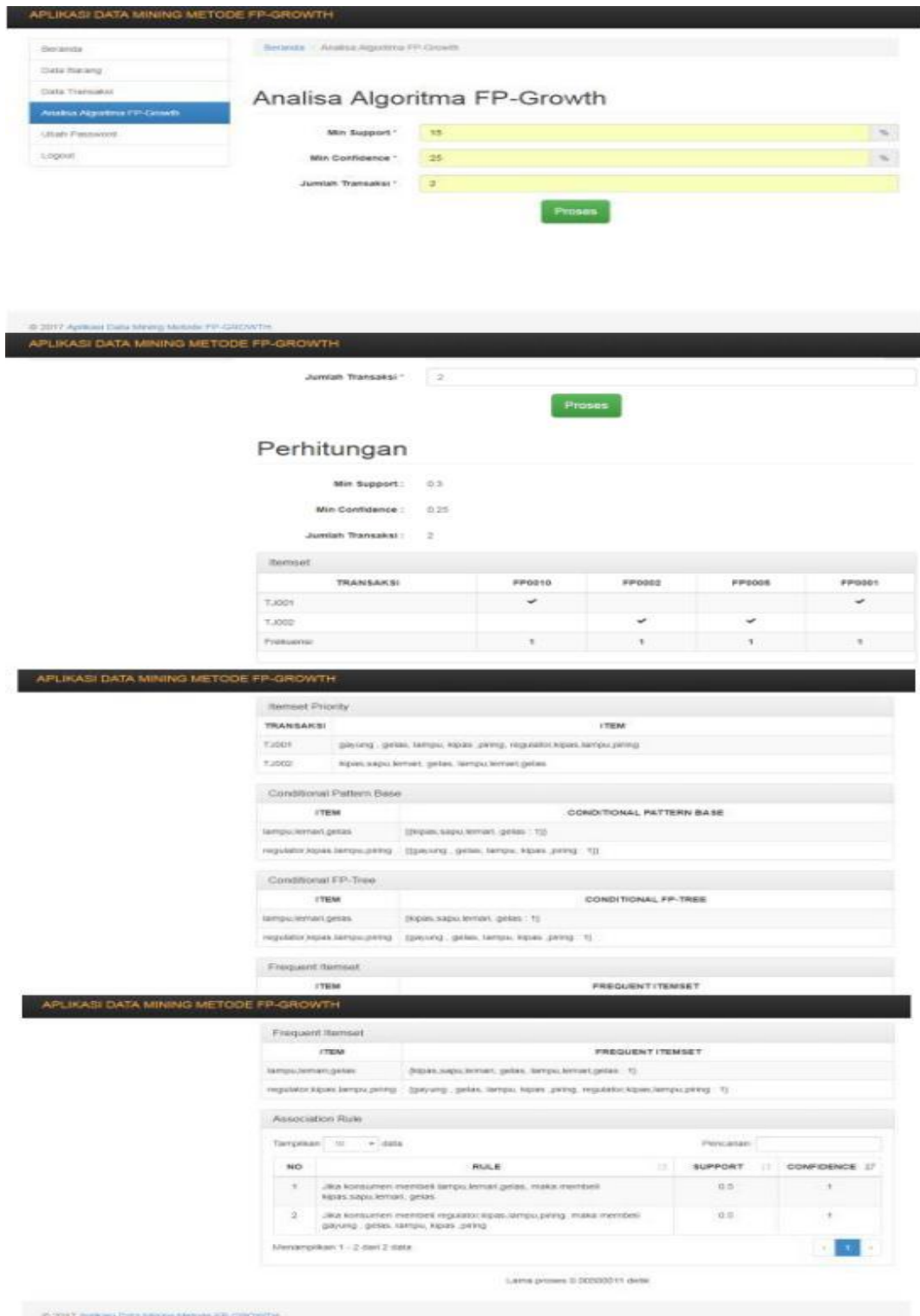
Gambar 2. Tampilan Data Admin



Gambar 3. Tampilan Data Produk



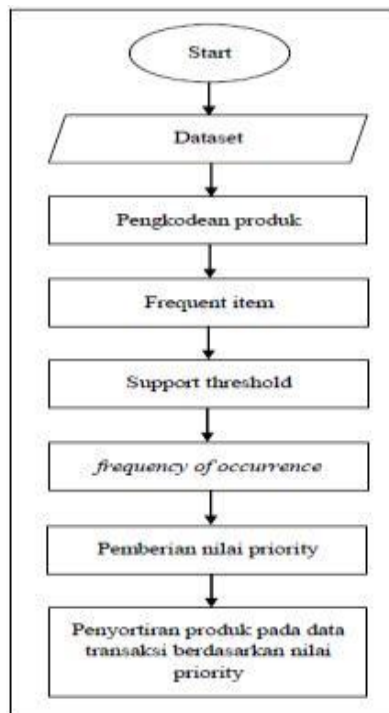
Gambar 4. Tampilan Analisis Algoritma FP-Growth



Gambar 5. Hasil Dari Analisa Algoritma FP-Growth

B. PEMBAHASAN

Proses *Data Mining* yang diterapkan pada penelitian ini menggunakan metode *Market Basket Analysis (Rule Association)* dengan *Algoritma Apriori*. Langkah – langkah dalam penerapan proses *Data Mining* yaitu pengumpulan data, *Inisialisasi* data, pembentukan *Association Rule*, dan pengambilan kesimpulan.



Gambar 6. Flowchart Sistem Kerja Dalam tahap pencarian *frequent*

itemset algoritma yang akan digunakan adalah *fp-growth*. Algoritma ini memiliki dasar pengetahuan mengenai *frequent itemset* yang telah diketahui sebelumnya untuk memproses informasi lebih lanjut. Dalam algoritma *Fp-growth* terdapat tiga tahapan penting, namun sebelum masuk pada tahapan untuk mempermudah pembentukan *frequent itemset* akan dilakukan pengkodean terhadap produk yang terdapat pada data yang sudah siap digunakan untuk proses *mining*, proses pengkodean ini berdasarkan inisial setiap produk yang terdapat pada transaksi. Berikut adalah pengkodean dari setiap produk dapat dilihat pada tabel dibawah ini:

Tabel 1. Hasil Pengkodean Produk

Produk	Kode Produk	Produk	Kode Produk
Televisi	TV	Tempat Sampah	TS
Kulkas	KU	Vint.co	VC
Antena TV	ANT	Hanger cjp	HC
K.A Maspion	KM	Ember	EB
Meja Medium	MM	Sapu ijuk	SI
Kipas Angin Miyako	KAM	Sikat Wc	SW
Rice Cooker	RC	Gelas	GL
Jam Dinding	JD	Bak Mandi	BM
Karpet	KRT	Stok Kontak 4 lubang	SK
Rak	RAK	Kain Pel Dragon	KPD
Kipas Angin Biasa	KAB		
Regulator	RGL		
Lampu	LMP		

Fp-tree adalah struktur data yang digunakan oleh algoritma *fp-growth* dalam penentuan *frequent itemset*. Kelebihan dari *Fp-tree* adalah hanya memerlukan dua kali pemindaian data transaksi yang terbukti sangat efisien. Data yang digunakan dalam tahapan ini adalah data penjualan yang telah dilakukan pengkodean pada setiap produknya untuk memudahkan saat pembuatan *tree* bisa dilihat pada tabel berikut :

Tabel 2. Penyederhanaan Data Penjualan

Kode	Produk	Kode	Produk
FP0001	TV, ANT	FP0077	RAK, GL
FP0012	TV, MM	FP0078	KPD, SW
FP0016	KIM, JD	FP0090	LMP,SK
FP0024	LMP, SK	FP0094	TS, SI
FP0032	RAK, TS, SI	FP0097	BM,KPD
FP0046	KAB, KRT	FP0104	KIM, RC
FP0047	EB, SW, BM	FP0113	KRT, SI
FP0054	KP, EB	FP0116	TV, SK
FP0055	LDG, OR, FC	FP0123	KU, SK
FP0056	LDG, OR, FC, HC	FP0130	SI, SW, KPD
FP0067	RC, SK	FP0141	RAK, RC
FP0070	KU, SK	FP0142	KAB, SK
FP0077	RAK, GL	FP0145	KIM, MM
FP0078	KPD, SW	FP0149	EB, SI, SW, KPD

Tahap selanjutnya adalah frekuensi kemunculan dari setiap produk dari data transaksi penjualan untuk melihat produk mana saja yang dapat di proses dalam tahap selanjutnya, seperti pada tabel dibawah ini:

Tabel 3. Frekuensi Kemunculan setiap Produk

Produk	Frekuensi	Produk	Frekuensi
KRT	2	HC	3
KAB	2	GL	4
LMP	4	VC	2
SK	3	KPD	3
RC	3	TS	2
SI	3	RGL	3
SW	5	KRT	2
EB	4	KAB	3
BM	3	LMP	2

Setelah frekuensi dari setiap produk diketahui maka selanjutnya adalah menentukan *minimum support*, seperti pada tabel berikut ini:

Tabel 4. Penghapusan Data Tidak frequent

Produk	Frekuensi
KRT	2
KAB	2
LMP	4
SK	3
RC	3
SI	3
SW	5
EB	4
BM	3
HC	3
GL	4
VC	2
KPD	3
TS	2
RGL	3
KRT	2
KAB	3
LMP	2

Setelah didapatkan data yang memenuhi *minimum support* selanjutnya akan dilakukan pengurutan produk berdasarkan frekuensi kemunculan terbanyak dan ditentukan *priority* untuk setiap produk, seperti pada tabel dibawah ini:

Tabel 5. Pengurutan Berdasarkan Frekuensi dan Penentuan Priority

Produk	Frekuensi	Priority
KRT	2	1
KAB	2	2
LMP	4	3
SK	3	4
RC	3	5
SI	3	6
SW	5	7
EB	4	8
BM	3	9
HC	3	10
GL	4	11
VC	2	12
KPD	3	13
TS	2	14
RGL	3	15
KRT	2	16
KAB	3	17
LMP	2	18

Setelah *priority* didapatkan, selanjutnya kemunculan produk akan dilakukan pengecekan setiap transaksi untuk diurutkan berdasarkan *priority* dari masing-masing produk. Berikut adalah hasil dari pengurutan data dapat dilihat pada tabel di bawah ini :

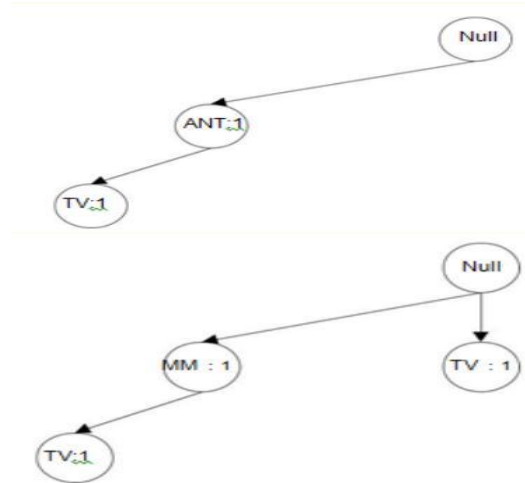
Tabel 6. Hasil Pengurutan Data

Kode	Produk(ordered item)	Kode	Produk(Ordered item)
FP0001	TV, ANT	FP0054	KP, EB
FP0012	TV, MM	FP0055	LDG, OR, FC
FP0016	KIM, JD	FP0056	LDG, OR, FC, HC
FP0024	LMP, SK	FP0067	RC, SK
FP0032	RAK, TS, SI	FP0070	KU, SK
FP0046	KAB, KRT	FP0077	RAK, GL
FP0047	EB, SW, BM	FP0078	KPD, SW
		FP0090	LMP, SK
		FP0094	TS, SI
		FP0097	BM, KPD
		FP0104	KIM, RC
		FP0113	KRT, SI
		FP0116	TV, SK
		FP0123	KU, SK
		FP0130	SI, SW, KPD
		FP0141	RAK, RC
		FP0142	KAB, SK
		FP0145	KIM
		FP0149	EB, SI, SW, KPD

Setelah mendapatkan data yang sesuai, selanjutnya setiap transaksi penjualan yang terdapat pada tabel 6 akan bangkitkan dengan struktur data *fp-tree*. Berikut adalah penerapan struktur data *fp-tree* : Transaksi dengan kode FP0001: TV, ANT. Diberikan *support count* 1 dan dua *node* dengan TV kebutuhan fungsional yang telah sebagai *parent* dan ANT sebagai *child* didefinisikan.

Tabel 8. Pengujian Pemilihan Data Transaksi

Kasus dan hasil uji (data normal)			
Data masukan	Yang diharapkan	Pengamatan	Kesimpulan
Data transaksi penjualan.xlsx	Dapat menampilkan alamat dari file yang dipilih	Alamat file muncul dalam <i>gridview</i>	[√] Diterima [] Ditolak
Kasus dan hasil uji (data salah)			
Data masukan	Yang diharapkan	Pengamatan	Kesimpulan
Alamat file kosong	Muncul pesan kesalahan agar memilih file terlebih dahulu	Menampilkan pesan kesalahan untuk memilih file terlebih dahulu	[√] Diterima [] Ditolak



Gambar 7 Struktur data fp-tree

Tahap implementasi sistem merupakan tahap untuk mengaplikasikan apa yang telah dirancang pada tahap perancangan sistem berdasarkan hasil analisis sistem. Implementasi sistem ini dilakukan sebagai wujud nyata hasil analisis dan perancangan sistem ke dalam bahasa pemrograman untuk diterapkan pada keadaan yang sesungguhnya. Hasil dari penerapan tersebut diharapkan dapat menjadi sebuah sistem yang siap diuji dan digunakan. Pengujian yang akan dilakukan adalah dengan menggunakan metode *white box* dan *black box* serta pengujian beta dengan mewawancarai pengguna yang akan menggunakan perangkat lunak ini.

Tabel 7. Rencana Pengujian

No	Item Uji	Detail Pengujian	Jenis Uji
1.	Algoritma <i>FP-Growth</i>	Uji Algoritma	<i>White box</i>
2.	Pemilihan Data Transaksi	Pengujian validasi Pemilihan Data Transaksi	<i>Black Box</i>
3.	Penyimpanan ke dalam database	Pengujian validasi Penyimpanan ke dalam database	<i>Black Box</i>
4.	<i>Preprocessing Data</i>	Pengujian validasi <i>Preprocessing Data</i>	<i>Black Box</i>
5.	Proses Asosiasi	Pengujian validasi proses asosiasi	<i>Black Box</i>

Pengujian *black box* dilakukan dengan tujuan untuk mengetahui apakah semua fungsi perangkat lunak telah berjalan semestinya sesuai dengan Tabel 8 di bawah ini merupakan *scenario* pengujian *Equivalence Class Partitioning* yang dilakukan pada bagian proses asosiasi.

Tabel 9. Pengujian Proses Asosiasi

Kasus dan hasil uji (data normal)			
Data masukan	Yang diharapkan	Pengamatan	kesimpulan
Nilai <i>minimum support</i> : 2 Nilai <i>minimum confidence</i> : 60%	menghasilkan aturan asosiasi yang sesuai dengan nilai <i>minimum support</i> dan <i>minimum confidence</i> dan proses data <i>mining association rule</i> .	Dapat menampilkan aturan asosiasi dengan ketentuan nilai <i>minimum support</i> dan nilai <i>minimum confidence</i> yang di-inputkan	[✓] Diterima [] Ditolak
Kasus dan hasil uji (data salah)			
Data masukan	Yang diharapkan	Pengamatan	kesimpulan
Nilai <i>minimum support</i> : abc nilai <i>minimum confidence</i> : abc	Muncul pesan kesalahan agar menyesuaikan masukan yang sesuai	Menampilkan pesan kesalahan agar menyesuaikan masukan yang sesuai	[✓] Diterima [] Ditolak

Pengujian ini menguji perangkat lunak yang telah dibangun apakah menghasilkan data yang diinginkan dan sesuai dengan hasil penerapan metode *association rule* dengan algoritma *FP-Growth*. Pengujian ini menggunakan 150 data transaksi penjualan pada tabel 2, setelah dilakukan *preprocessing* jumlah data menjadi 26 transaksi lalu di *generate* dengan nilai *minimum support* ditentukan sebesar 2 dan nilai *minimum confidence* 60%. Data hasil *preprocessing* dapat dilihat pada tabel D-3 dalam lampiran D dan data *rule* yang dihasilkan dari penerapan metode *association rule*.

4. KESIMPULAN DAN SARAN

Berdasarkan pada hasil perancangan, implementasi, dan pengujian terhadap “Implementasi Data Mining Pada Penjualan Produk Di Toko Cahaya Setya Menggunakan *Algoritma Fp-Growth*”, dapat disimpulkan sebagai berikut:

1. Bahwa semakin kecil nilai *minimum support* maka akan semakin banyak aturan asosiasi yang dihasilkan, dan sebaliknya semakin besar nilai *minimum support* maka akan semakin sedikit aturan asosiasi yang dihasilkan. Semakin kecil nilai *minimum confidence*, kemungkinan nilai kepastian aturan asosiasi yang dihasilkan akan semakin banyak dan semakin besar nilai *minimum confidence* maka akan semakin sedikit aturan asosiasi yang dihasilkan.
2. Metode *Association Rule* dengan menggunakan Algoritma *FP-Growth* yang dibangun melalui Aplikasi dapat membantu Toko Cahaya Setya
3. Diharapkan penelitian ini dapat menjadi referensi bagi penelitian selanjutnya dengan mengembangkan aplikasi dengan menggunakan algoritma yang lainnya.

DAFTAR PUSTAKA

- [1] Erwin. (2015). Analisis Market Basket dengan Algoritma Apriori & *FP-Growth*. Jurnal Generic No.26-30.
- [2] Jiawei, H., Kamber, M. (2015). *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers.
- [3] Fajrin, Alfannisa Annurullah dan Algifanri Maulana. (2018). *Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor*. Jurnal Ilmu Komputer (KLIK) Teknik Informatika, Universitas Putera Batam Volume 05, No.01 Februari 2018.
- [4] Firdaus, Diky. 2017. *Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer*. Jurnal Format Volume 6 Nomor 2 Tahun 2017
- [5] David Samuel, penerapan Stuktur *FP-Tree* dan Algoritma *FP-Growth* dalam Optimasi Penentuan *FrequentItemset*, Institut Teknologi Bandung, Vol. 1, 2008.
- [6] Saad, A. dan Alghamdi, A., (2016). Efficient Implementation of *FP Growth* Algorithm-Data Mining on Medical Data. 11(12), pp.7–16.
- [7] Cahyana, Nur Heri et all. (2017), Aplikasi Penerimaan Siswa Baru Berbasis WEB (SMK Negeri 3 Yogyakarta). *TELEMATIKA* Vol. 10, No. 1, JULI 2013: 1 – 8.
- [8] Sijabat, ALimancon,(2015). Penerapan data mining untuk pengolahan data siswa dengan menggunakan metode *decision tree*. Jurnal Informasi dan teknologi Ilmiah Volume 5 No. 3 ISSN:2339-210x
- [9] Ellis, software Engineering, *SIGSoft Software Engineering Notes* Vo. 32 (2015)
- [10] Web Dalam Perkembangan and Winda Febriani Kusuma (2015). Pengembangan halaman web, menggunakan XML Dalam Perkembangan WEB 2.0, Jurnal Teknik Informatika 6 (2):8