

## MENCARI POLA PEMBELIAN KONSUMEN MENGGUNAKAN ALGORITMA FP-GROWTH

Arandia Raka Wijaya<sup>1</sup>, Arief Jananto<sup>2</sup>

<sup>1,2</sup>Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Stikubank  
e-mail: <sup>1</sup>arnandiawijaya@gmail.com, <sup>2</sup>ajananto09@gmail.com

### ABSTRAK

Data Mining adalah proses pencarian informasi dengan melakukan penggalian dari pola pola data transaksi dengan tujuan dapat memperoleh sebuah informasi yang berharga untuk mengelola data yang sangat besar. Pada Toko Kita setiap harinya terjadi ratusan transaksi. Dalam mempermudah melakukan pengelolaan penyediaan barang maka perlu diketahui produk apa saja dan itemset antar kaitan barangnya untuk membantu menangani penyediaan barang. Dalam kaitannya dalam mengatasi permasalahan penyediaan yang ada pada Toko Kita, data mining memiliki beberapa teknik salah satunya adalah teknik assosiasi. Teknik assosiasi dilakukan untuk menemukan aturan assosiatif antara suatu kombinasi item. Sedangkan beberapa metode algoritma yang dapat digunakan salah satunya adalah algoritma FP-GROWTH yaitu dalam pembangunan frequent itemsetnya algoritma FP-GROWTH menggunakan struktur data tree atau disebut FP-Tree, dari struktur FP-tree ini nantinya frequent itemset dapat langsung diekstrak dan diketahui hasilnya. Adapun dalam penelitian ini data yang digunakan adalah data transaksi pada Toko Kita dari rentang waktu bulan November 2016 sampai Oktober 2017. Dimana dalam penelitiannya dihasilkan beberapa aturan assosiasi salah satunya yaitu  $\{item1=899898910012GG\ FILTER\ 12\} \Rightarrow \{frequent=899190610101DJARUM\ SUPER\ 12\}$  yang berarti item GG FILTER 12 muncul bersamaan dengan item DJARUM SUPER 12 dengan tingkat dukungan 0.02 dan tingkat kepercayaan 1.00.

**Keyword:** Data Mining, Association Rule, FP-GROWTH

### 1. PENDAHULUAN

Data Mining adalah proses pencarian informasi dengan melakukan penggalian dari pola pola data transaksi dengan tujuan dapat memperoleh sebuah informasi yang berharga. Data mining sangat perlu dilakukan terutama untuk mengelola data yang sangat besar untuk memudahkan aktifitas recording suatu transaksi agar dapat memberikan informasi yang tepat bagi para penggunanya. Pada Toko Kita setiap harinya terjadi ratusan transaksi. Dalam mempermudah melakukan pengelolaan penyediaan barang maka perlu diketahui produk apa saja dan itemset antar kaitan barangnya untuk membantu menangani penyediaan barang yang dihadapi pemilik. Untuk itu semakin bertambahnya jumlah data setiap harinya maka peran analisis secara manual perlu digantikan dengan aplikasi berbasis komputer sehingga proses analisis dapat lebih mudah dan cepat.

Dalam kaitannya dalam mengatasi permasalahan penyediaan yang ada pada Toko Kita, data mining memiliki beberapa teknik salah satunya adalah teknik asosiasi. Teknik asosiasi dilakukan untuk menemukan aturan assosiatif antara suatu kombinasi item. Aturan assosiasi meliputi dua tahap yaitu mencari kombinasi yang paling sering terjadi dari suatu itemset dan mendefinisikan kondisi dan hasil (untuk tergantung aturan asosiasi). Adapun dalam penelitian ini akan dilakukan pembahasan bagaimana mengimplementasikan salah satu algoritma dalam data mining, yaitu algoritma FP-Growth. Algoritma ini merupakan alternatif yang dapat digunakan untuk menentukan himpunan data yang sering muncul dalam sekumpulan data. Algoritma ini merupakan bagian dari teknik asosiasi datamining.

Diharapkan dalam penelitian ini nantinya, setelah dilakukan implementasi data mining dari data transaksi didapatkan sebuah informasi atau knowledge berupa pola transaksi yang dihasilkan sehingga dapat mendukung kemudahan penyediaan inventory secara lebih efektif pada Toko Kita.

### 2. TINJAUAN PUSTAKA

[2]Penelitian ini membahas bagaimana cara memaksimalkan penjualan pada supermarket dengan mengetahui pola pembelian konsumen menggunakan data mining. Adanya data yang menumpuk dan berjumlah sangat banyak merupakan alasan utama digunakannya algoritma apriori untuk menunjang pengambilan keputusan. Peneliti juga ingin membandingkan perbedaan pembelian konsumen antar

cabang lainnya. Perbedaan letak dalam pemilihan cabang tentunya akan berbeda pula pola pembelian konsumennya, untuk itu hal ini perlu dianalisis lebih jauh lagi sehingga dapat menghasilkan hasil yang lebih maksimal dan keuntungan yang diperoleh.

[3]Melakukan sebuah penelitian yaitu untuk membahas bagaimana melakukan sebuah promosi yang tepat. Penelitian yang dilakukan yaitu melakukan proses pengolahan data mining yang diambil dari data STMIK TRIGUNA DHARMA 1 berupa asal sekolah, daerah asal, jurusan, pilihan prodi untuk melakukan proses promosi perguruan tinggi. Dari data yang diolah akan dilakukan proses pengambilan keputusan dalam melakukan sasaran promosi yang tepat dan potensial sehingga dapat meminimalisir waktu dan biaya yang diperlukan. Adapun penelitian ini akan menggunakan salah satu teknik algoritma dari data mining yaitu FP-GROWTH. Algoritma ini merupakan bagian dari teknik asosiasi yang dapat menemukan pola kecenderungan data yang paling sering muncul dalam sekumpulan data.

[4]Membahas bagaimana mengatasi masalah persediaan barang pada Ahas Rahmat Motor Blora. Masalah yang dialami pada dealer ini yaitu pada saat suku cadang tertentu mengalami kekosongan maka suku cadang yg lain juga ikut dibeli. Hal ini tentu menjadi kendala dikarenakan banyaknya jenis dari suku cadang sepeda motor. Dengan adanya masalah tersebut penulis berusaha mengatasi dengan melakukan pengolahan data mining dari data penjualan suku cadang. Market Basket Analysis atau yang dapat disebut dengan teknik asosiasi merupakan salah satu teknik dari data mining, dimana teknik ini dapat digunakan untuk mencari pola kecenderungan pembelian konsumen antara produk satu dengan kaitan pembelian bersamaan produk lainnya. Diharapkan setelah dilakukan penelitian ini, penulis dapat mendapatkan hasil informasi atau knowledge, yang nantinya dapat digunakan untuk membantu mengatasi permasalahan pengiriman suku cadang.

Dari ketiga penelitian diatas dapat disimpulkan bahwa penggunaan data mining dapat digunakan untuk pencarian informasi yang berguna dalam berbagai bidang. Berbagai algoritma juga dapat diterapkan untuk proses asosiasi mining, salah satunya yaitu FP- GROWTH yang merupakan tingkat lanjut penyempurnaan dari algoritma Apriori.

## 2.1 Assosiasi

Dalam penentuannya aturan asosiasi menggunakan 2 buah parameter atau patokan penghitungan yaitu support dan confidence. Support menunjukkan persentasi jumlah transaksi yang berisi X dan Y. Sedangkan Confidence menunjukkan persentasi banyaknya Y pada transaksi yang mengandung X. Bentuk persamaan matematisnya dapat dituliskan seperti ini[5]

$$\text{Support}(X,Y) = \frac{\text{jumlah transaksi mengandung } X \text{ dan } Y}{\text{total transaksi}}$$

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{jumlah transaksi mengandung } X \text{ dan } Y}{\text{jumlah transaksi mengandung } X}$$

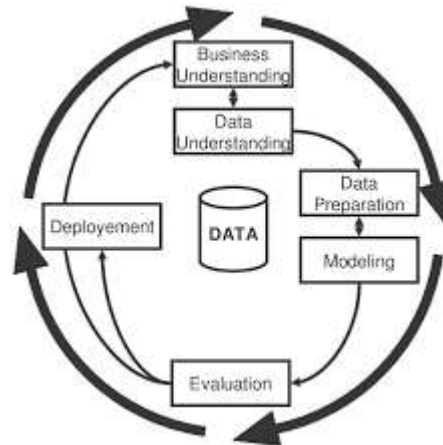
## 2.2 Algoritma FP-GROWTH

Dalam pencarian frequent itemset algoritma FP-GROWTH menggunakan struktur data tree atau yang sering disebut FP-Tree, dari struktur FP-tree inilah nantinya frequent itemset dapat langsung diekstrak dan diketahui hasilnya. Untuk pengaplikasiannya FP-GROWTH dibagi menjadi 3 tahapan utama yaitu sebagai berikut[6]:

1. Tahap Pembangkitan conditional pattern base
2. Tahap Pembangkitan conditional FP-Tree
3. Tahap pencarian frequent itemset

## 3. METODE PENELITIAN

Cross-Industry Standard Process for Data Mining (CRISP-DM) secara umum menjelaskan proses pengembangan data mining dalam enam tahap, yang telah ditetapkan pada tahun 1996 sebagai standar untuk proses analisis. Enam fase dalam CRISP-DM yaitu[1]

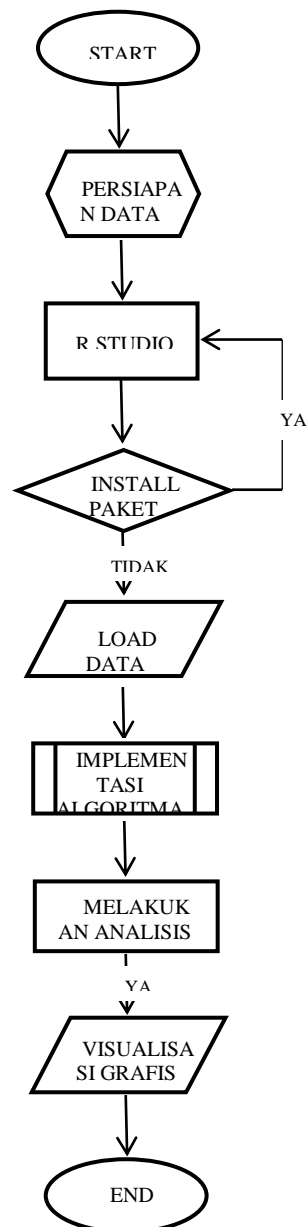


**Gambar 1. Siklus CRISP-DM**

1. Business Understanding  
Memahami tujuan dan kebutuhan kemudian menterjemahkan pengetahuan ke dalam pendefinisian masalah dalam penggunaannya untuk pemrosesan data mining.
2. Data Understanding  
Dimulai dengan pengumpulan data kemudian dilanjutkan untuk mendapatkan pemahaman tentang data, apa saja yang dibutuhkan untuk bagaimana data akan diproses dan diolah nantinya.
3. Data Preparation  
Meliputi semua kegiatan dalam membangun dataset dari awal sampai akhir mencakup pemilihan record, atribut data, serta proses pembersihan bagian yang tidak diperlukan.
4. Modelling  
Tahap ini dilakukan pemilihan serta penerapan teknik pemodelan yang disesuaikan untuk mendapatkan hasil yang optimal
5. Evaluation  
Dilakukan evaluasi terhadap keefektifan model digunakan apakah model yang digunakan telah memenuhi tujuan yang ditetapkan pada fase sebelumnya
6. Deployment  
Pengetahuan serta informasi yang sudah didapatkan akan dipresentasikan kedalam bentuk khusus yang lebih mudah dipahami oleh pengguna nantinya

### 3.1 Flowchart Implementasi Data Mining

Dalam pemrosesannya tools yang digunakan adalah R studio. Proses yang dilakukan yaitu melakukan implementasi data mining untuk mendapatkan informasi sesuai dengan tujuan yang telah dibahas sebelumnya. Pada tahap ini penggalian informasi dilakukan dengan menggunakan salah satu teknik data mining yaitu *association rule*.



**Gambar2. Flowchart implementasi data mining**

#### 4. HASIL DAN PEMBAHASAN

Dalam pembahasannya, peneliti menggunakan tools R studio untuk melakukan proses pengolahan dalam data mining. Berikut langkah-langkah untuk mendapatkan hasil frequent itemset, yaitu:

1. Install Packages  
Dalam proses penggalian frequent itemset menggunakan algoritma FP-GROWTH diperlukan 2 packages yang harus diinstall yaitu packages arulesViz dan packages rCBA.
2. Import Data  
Berikut gambar data yang digunakan:

	A	B	C	D	E	F	G	H
1	item1	item2	item3	item4	item5	item6	item7	item8
2	89927757121	GIGI	OPUS GULA	GULA PASIR				
3	8999999531	08968640172	89988881481	899888814	899888812	8999999028230		SAM SOE KRETEK
4	08968640170	89910021054	89910021039	SAMBAL A	89988881001200	G FILTER 12		
5	8998888100	89918881111	89999992940	899600157	899600130	899275301	899100210548	KAPAL
6	8991906101300	JARUM SUPER MLD 2						
7	8999999076	89970218704	89923040367	89970131488	SALEP BIRI			
8	8992702003	89920964044	BEAR BRAND	189ML				
9	899678871153	KWETIAU KUAH	AYM B					
10	8998888110	89988885201	89988881001500	SIGNAL	899888810	899888866	899294631	899281281
11	8999999000405	SAMPICERNA	KRETEK 1					
12	8998888101094	QUA HTL	1500ML					
13	89988886803	899888868023AN	LVO SBN	CARING 80				
14	89932263021	GULA JW 0.25G	GULA JAWA	1/4KG				
15	RAPIKA SAC	89927270052	BATTACK	JAZI 5MRBK				
16	BAN100	TALE ALE	ALE ALE	JWA/STR/AG				
17	ALE ALE	A GULA ASEM	89960016002	899276100	899600160	899277231	899888866	899100210
18	GAS 3KG	89910023072	BAN100	TEH BANDULAN	3AKI			
19	TERIGU 25	89988881101	89932283025	PLASTIK	89941711	899100210	89910021151	AKA GRA
20	PINUS	PI POP ICE	PC	89988882003	GULA 0.25	TIGA	TEBU	SARI MANIS 1
21	8998888100	89988881101	600G	SURYA 10				
22	89928261131	TELUH	TELUH					
23	8998888120	TELUH	TELUH					

Gambar3. Data transaksi yang digunakan

1. Mengubah Tipe Data  
Data transaksi terlebih dahulu dirubah ke dalam format “transactions” agar data dapat diproses. Berikut dibawah gambar hasil bentuk data setelah menjadi format “transactions”.

```

> inspect(trans)
      items                                     transactionID
[1]  {item1=899277571262MYTEA TEH OLONG 35,
      item2=GIGI OPLOSAN GIGI,
      item3=GULA GULA PASIR,
      frequent=8991906101010JARUM SUPER 12}          1
[2]  {item1=899999903320LIFEBUOY SMOPO ANT,
      item2=089686401721INDOFODD SAUS TOMA,
      item3=899888814818PEGL LINA KOMPLIT,
      item4=899888814918KUKU BIMA KOMPLIT,
      item5=899888813018TOLAK ANGIN KOMPLI,
      item6=8999999028230JI SAM SOE KRETEK,
      frequent=8991906101010JARUM SUPER 12}          2
[3]  {item1=089686017076SARIMI RBS AYM BWG,
      item2=899100210548KAPAL API SPECIAL,
  
```

Gambar4. Data transaksi setelah dirubah format

2. Menampilkan frequent itemset  
Dengan uji support minimal 0.02 dan dengan consequent DJARUM SUPER12 didapatkan hasil sebagai berikut:

```

> inspect(rules)
      lhs                                     rhs
t confidence lift
[1] {item1=GAS 3KG GAS ELPIJI 3KG} => {frequent=8991906101010D.
0 1 1
[2] {item1=89988881001200 FILTER 12} => {frequent=8991906101010D.
8 1 1
[3] {item1=TERIGU TERIGU TEGU} => {frequent=8991906101010D.
8 1 1
[4] {item1=GULA GULA PASIR} => {frequent=8991906101010D.
8 1 1
[5] {} => {frequent=8991906101010D.
0 1 1
>
  
```

Gambar 5. Hasil frequent itemset

Pada gambar diatas dapat dijelaskan yaitu:

- [1] Jika seseorang melakukan pembelian GAS LPG 3kg maka orang tersebut juga melakukan pembelian DJARUM SUPER 12 dengan nilai support 0.02 dan nilai confidence 1.00
- [2] Jika seseorang melakukan pembelian GG FILTER 12 maka orang tersebut juga melakukan pembelian DJARUM SUPER 12 dengan nilai support 0.02 dan nilai confidence 1.00
- [3] Jika seseorang melakukan pembelian TERIGU maka orang tersebut juga melakukan pembelian DJARUM SUPER 12 dengan nilai support 0.02 dan nilai confidence 1.00
- [4] Jika seseorang melakukan pembelian GULA PASIR maka orang tersebut juga melakukan pembelian DJARUM SUPER 12 dengan nilai support 0.02 dan nilai confidence 1.00

#### 5. Hasil Graph

Berikut hasil Grafik yang disajikan dari hasil frequent itemset yang didapatkan:



**Gambar6. Hasil Grafik**

## 5. KESIMPULAN

Berikut kesimpulan yang dapat penulis sampaikan setelah dilakukan penelitian tentang Implementasi Data mining dengan menggunakan algoritma FP-GROWTH study kasus Toko Kita:

1. Pemrosesan Data mining menggunakan algoritma FP-GROWTH berhasil diimplementasikan untuk mendapatkan analisis asosiasi market basket di Toko Kita. Hal ini terlihat peneliti berhasil mendapatkan item apa saja yang saling muncul secara bersamaan dalam kumpulan data transaksi yang terjadi antara bulan November 2016 – Oktober 2017

2. Dengan hasil penelitian yang telah dilakukan ini kita dapat mengetahui pasangan antar barang yang paling sering dibeli secara bersamaan dengan item DJARUM SUPER 12 yaitu item GAS LPG3kg, GG FILTER 12, TERIGU, dan GULA PASIR dengan masing masing nilai dukungan yaitu sebanyak 0.02 dan nilai kepercayaan 1.00

3. Dari hasil frequent itemset yang sudah didapatkan maka pihak pemilik Toko Kita dapat melakukan persediaan barang lebih mudah.

Saran yang dapat peneliti sampaikan untuk lebih lanjut adalah:

1. Agar mendapatkan hasil yang lebih baik sumber data yang digunakan dapat lebih banyak dan lebih lengkap sehingga rule asosiasi yang terbentuk akan lebih beragam.
2. Untuk pengembangan lebih lanjut, sebaiknya aplikasi dapat segera diterapkan pada Toko Kita.

## DAFTAR PUSTAKA

- [1] Shearer, Colin (2000) *The Crisp-DM Model: The New Blueprint for Data Mining*. Journal of Data Warehousing, Vol. 5, No. 4, p. 13-22.
- [2] Failasufa, Fusna. (2014) Penerapan Data Mining untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Apriori Pada Data Transaksi Penjualan, *Skripsi*, Program Studi Teknik Informatika, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta.

- [3] Ikhwan, Ali, Dicky Nofriansyah dan Sriani (2015) Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi PendidIndomie, *Jurnal Ilmiah Sains dan Komputer*, 3 (14), pp. 2.
- [4] Nugroho, Tri S. (2015) Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Aturan Asosiasi Pada Penjualan Suku Cadang Sepeda Motor Di Ahass Rahmat Motor Blora, *Skripsi*, Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Semarang.
- [5] Prasetyo, Philips Kokoh (2006) Itemset, Support, dan Confidence, Website: <https://philips.wordpress.com/2006/05/18/itemset-support-dan-confidence/>
- [6] Ardani, Nur R dan Nur Fitriana (2016) Sistem Rekomendasi Pemesanan Sparepart Dengan Algoritma FP-GROWTH, *Jurnal STMIK AMIKOM*, No.3,3, 6-7 Februari, p. 99.