

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Teori Umum**

##### **2.1.1 Pengertian Ritel**

Dalam penelitian (Nurhayati & Yuanita 2015) bidang usaha *retail modern* umumnya adalah suatu aksi berupaya menjual aneka barang atau jasa dagangannya sebagai konsumsi langsung maupun tidak langsung. Dalam rantai perdagangan bidang usaha ritel merupakan elemen akhir dari proses distribusi barang/jasa dan berkomunikasi langsung dengan konsumen.

Menurut (Kurniawan, Gata, and Wiyana 2018) mengatakan dalam mendirikan usaha ritel, perusahaan kosmetik tidak terlepas dalam penggunaan teknologi informasi yang mendukung kecakapan dalam penjualan produk-produk. Penggunaan teknologi informasi menjadi suatu kewajiban saat ini, melihat pesat dan ketatnya persaingan bisnis di bidang tersebut.

Sebab maraknya usaha ritel penjualan produk kosmetik, maka secara otomatis persaingan bidang usaha tersebut akan semakin ketat. Setiap perusahaan akan terus berinovasi dalam pelayanan pemasaran yang terbaik diimbangi dengan peningkatan kualitas produk yang diciptakan.

#### **2.2 Teori Khusus**

##### **2.2.1 Knowledge Discovery in Database (KDD)**

*Knowledge Discovery in Database* (KDD) diartikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari himpunan data. Proses *knowledge discovery* menyertakan hasil dari proses *data mining* (proses ekstrak

kecondongan suatu pola data), lalu mengubah hasilnya menjadi informasi yang akurat dan mudah dipahami (Fajrin et al. 2018). KDD juga didefinisikan sebagai keseluruhan proses *non-trivial* untuk menemukan dan menelaah pola suatu data, dimana pola yang ditemukan bersifat sah, baru, bermanfaat dan dapat dimengerti (Pohan et al. 2020)

Secara umum proses *Knowledge Discovery in Databases* terdiri dari (Elisa 2018):

1. *Data Cleaning*

Proses membuang *noise* dari data yang tidak konsisten.

2. *Data Integration*

Kombinasi data dari berbagai basis data ke dalam satu basis data yang baru.

3. *Data Selection*

Proses seleksi data yang berhubungan yang berasal dari basis data.

4. *Data Transformation*

Data diubah ke dalam format yang cocok dalam proses *mining*.

5. *Data Mining*

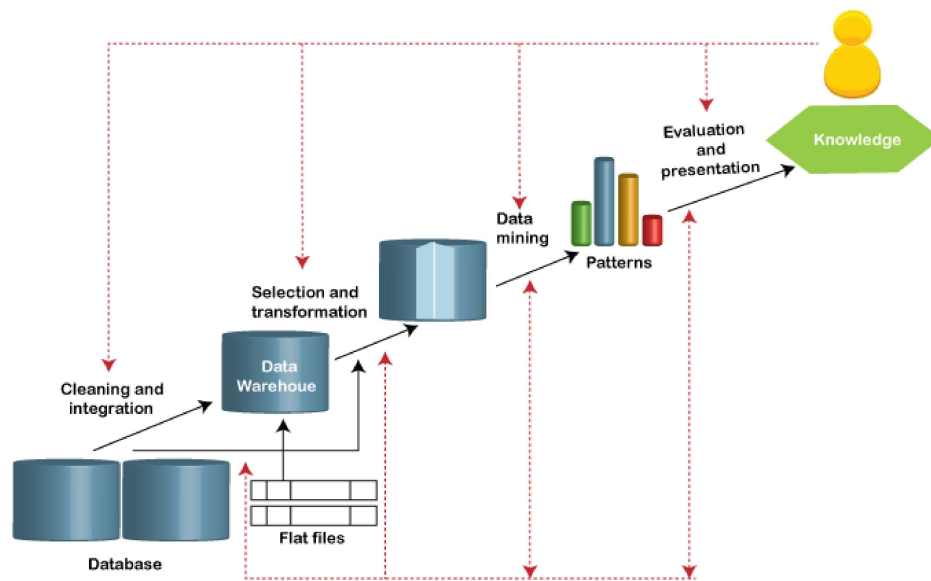
Metode yang diaplikasikan untuk mencari wawasan baru yang tersembunyi dari data.

6. *Pattern Evaluation*

Menganalisa *patterns* untuk didemonstrasikan ke dalam *knowledge based*.

7. *Knowledge Presentation*

Wujud dan presentasi pengetahuan mengenai teknik yang dipakai untuk mendapat wawasan baru yang diperoleh oleh *user*.



**Gambar 2.1** Proses KDD (Knowledge Discovery in Database)

### 2.2.2 Data Mining

Jurnal Penelitian (Junaidi 2019) menuliskan bahwa menurut Gartner group penggalian data merupakan metode mencari asosiasi baru yang mempunyai makna, pola dan tren sejumlah besar data yang tersimpan dalam gudang data, memakai pola teknologi informasi bersamaan dengan tekhnik matematika dan statistik. *Data mining* berpotensi menemukan korelasi antara item satu dengan item lainnya pada himpunan data transaksi yang besar. Berkat data yang selalu dikumpulkan dan disimpannya data dalam jangka yang berkepanjangan banyak industri terdorong dengan proses *data mining* dan *database*.

Teknik pada data mining dibagi menjadi beberapa kelompok, yaitu (Rasyidah and Dwiasnati 2021) :

1. Deskripsi yaitu menemukan karakteristik penting untuk menggambarkan pola data.

2. Klasifikasi yaitu menemukan kelompok untuk mendeskripsikan konsep data.
3. Estimasi mirip dengan klasifikasi. Namun, teknik estimasi menggunakan variabel numerik.
4. Prediksi untuk menemukan pola data memakai variabel untuk memperkirakan variabel lain.
5. Klastering untuk mengelompokkan data berdasarkan kesamaan dari objek yang sudah tampak.
6. Asosiasi adalah teknik *data mining* yang bermanfaat untuk mencari item yang timbul dalam waktu yang bersamaan.

### 2.2.3 Association Rule

Menurut (Wibowo and Jananto 2020) Analisis asosiasi atau *association analyze* menjadi salah satu cara *data mining* untuk pencarian aturan asosiatif pada gabungan item atau korelasi antar atribut. Analisis aturan asosiasi berguna untuk menemukan hubungan antar item dalam suatu data yang besar. Ikatan antar item yang ditemui akan disampaikan dalam bentuk aturan asosiasi atau itemset yang sering timbul.

Proses penggalian *association rules* adalah agar menemukan korelasi antar item dalam basis data. Dimulai dengan menggali *frequent itemset*, yaitu kombinasi yang paling sering muncul dalam sekelompok itemset dan wajib mencukupi *minimum support*. Pemeriksaan itemset yang sering timbul dari basis data transaksional yang besar merupakan hal yang paling sulit dalam *data mining*. Kenyataannya, data tidak diekstraks dari asal data tunggal namun berasal dari data yang tersalurkan dan heterogen. Wawasan yang baru diharapkan bisa membantu

jadwal kegiatan usaha yang lebih baik. Hal ini dinyatakan oleh (Kurniawan et al. 2018) dalam penelitiannya.

Menurut (Astrina et al. 2019) terdapat dua ukuran dalam analisis asosiasi yaitu *Support* dan *Confidence*.

- a. *Support* adalah pengukuran untuk menentukan seberapa besar item yang lebih mendominasi item dari keseluruhan transaksi.
- b. *Confidence* adalah pengukuran untuk menentukan korelasi antar dua item atau lebih berdasarkan suatu kondisi.

#### **2.2.4 Algoritma Frequent Pattern Growth (FP-Growth)**

Menurut (Simanjuntak and Windarto 2020) Algoritma *FP-Growth* adalah salah satu opsi yang bisa dipakai untuk menetapkan tumpukan data yang paling sering timbul (*frequent itemset*) dalam sebuah himpunan data. Pada algoritma *FP-Growth* memakai konsep penyusunan *tree*, yang biasa disebut *FP-Tree* dalam pencarian *frequent itemset* karena prinsip tersebut algoritma *FP-Growth* lebih efektif daripada algoritma Apriori.

Menurut (Elisa and Azwanti 2019) algoritma apriori lebih populer daripada algoritma *FP-Growth*. Oleh sebab itu, *weakness* dari algoritma Apriori dikembangkan oleh algoritma *FP-Growth*. *Frequent Pattern Growth (FP-Growth)* adalah salah satu opsi dari algoritma lain yang biasa digunakan untuk menunjukkan itemset yang sering dalam suatu set data. Dengan algoritma Apriori, akan mewujudkan kandidat untuk memperoleh itemset yang sering timbul. Penggunaan rancangan algoritma *FP-Tree* ini sangat efisien dari pendahulunya yaitu apriori.

Untuk memutuskan *frequent* itemset pada data transaksi tersebut, bisa diterapkan tahap-tahap berikut ini (Nastuti and Harahap 2019) :

1. Menentukan *minimum support*.
2. Menentukan *header frequent itemset*
3. Membuat *FP-Tree*
4. Membuat *conditional pattern* berdasarkan *FP-Tree*
5. Menentukan *frequent itemset*

### **2.2.5 Software RapidMiner Studio**

Mengutip dalam jurnal (Firmansyah and Eriswandi 2020) *Rapidminer* adalah *software* yang diciptakan oleh Dr. Markus Hofmann dari *Institute of Technology*

Blanchardstown dan Ralf Klinkenberg dari rapid-i.com dengan tampilan GUI (*Graphical User Interface*) untuk menyederhanakan penggunaan aplikasi ini. Aplikasi ini bersifat *open source* serta mudah digunakan di sistem operasi manapun. Dalam penggunaannya, kemampuan koding khusus tidak terlalu dibutuhkan sebab seluruh fasilitas sudah disediakan. *Rapidminer* diperuntukkan dalam *data mining*. Banyak metode yang disediakan oleh *rapidminer* mulai dari klasifikasi, klustering, asosiasi dan sebagainya.

### 2.2.6 Penelitian Terdahulu

Penelitian terdahulu dari penelitian ini adalah sebagai berikut :

**Tabel 2.1** Penelitian Yang Sudah Dilakukan Peneliti Sebelumnya

No.	Judul Penelitian	Peneliti & ISSN	Hasil & Pembahasan
1	Implementasi <i>Association Rule &amp; Frequent Pattern Growth</i> Untuk Penentuan Sistem Rekomendasi Keputusan Pembelian	Siti Hasuna, Widi Hastomo, Ellya Sestri dan Nawang Kalbuana (2020) No. ISSN 2581-2327	Dalam penelitian tersebut menghasilkan rekomendasi sistem yang membantu sebuah toko ritel di stasiun Bogor untuk menjual produk dengan kombinasi produk dengan <i>minimum support</i> 0.2 dan <i>minimimun confidence</i> 0.5 yaitu makanan pokok (nasi putih dengan lauk cepat saji yaitu sup buntut dan soto kari daging/ kaki sapi sebagai 2 kombinasi

			dengan nilai <i>confidence</i> yang paling tinggi yaitu 0.951 & 0.944.
2	Analisis Algoritma <i>FP-Growth</i> Untuk Rekomendasi Produk Pada Data <i>Retail</i> Penjualan Produk Kosmetik (Studi Kasus : Mt Shop Kelapa Gading)	Sigit Kurniawan, Windu Gata, Hari Wiyana (2018) No. ISSN 2089-9815	Pengelolaan data transaksi penjualan selama 1 tahun pada Mt <i>Shop</i> Kelapa Gading yang diubah menjadi itemset dan diolah memakai <i>Rapidminer</i> 8.0 ditemukan nilai <i>confidence</i> tertinggi adalah 89% dengan aturan tiap penjualan produk Masker Beras Putih dapat ditetapkan akan membeli Putih Langsung <i>Facial Foam</i> . Permasalahan yang ada di MT Shop Kelapa



			Gading, seperti menentukan posisi produk-produk yang ada di etalase penjualan dan merekomendasikan produk-produk berdasarkan dari hasil pengolahan data tersebut dengan hasil <i>confidence</i> mulai dari 50% sampai dengan 89%.
3	Teknik Data Mining Untuk Penentuan Paket Hemat Sembako Dan Kebutuhan Harian Dengan Menggunakan Algoritma <i>FP-Growth</i> (Studi Kasus Di Ulfamart Lubuk Alung)	Amelia Nastuti, Syaiful Zuhri Harahap (2019) No. ISSN 2615-1855	Peneliti menyimpulkan bahwa Algoritma <i>FP-Growth</i> memanfaatkan himpunan data <i>FP-Tree</i> , mengamati basis data dua kali, tidak menghasilkan item pilihan dalam proses <i>mining</i> , dan

			sangat efisien untuk <i>data mining</i> .
4	Penerapan <i>Data Mining</i> Untuk Merekomendasikan Item Berdasarkan Pola Asosiasi Data Penjualan Pada KI-KHA <i>SHOP</i> Menggunakan Algoritma <i>FP-Growth</i>	Hiwilma Cleta Ermanti, Lisna Zahrotun, S.T., M.Cs (2016) No. ISSN 2252-5459	Dari hasil <i>rules</i> tersebut didapatkan <i>lift ratio</i> tertinggi yaitu 204,081 dengan beberapa kemungkinan yaitu ketika membeli tas ransel SD hellokity maka item yang di rekomendasikan yaitu Wahana.
5	Analisa <i>Data Mining</i> Menggunakan <i>Frequent Pattern Growth</i> pada Data Transaksi Penjualan PT. Mora Telematika Indonesia untuk Rekomendasi Strategi Pemasaran	Harpa Erasmus Simanjuntak, Windarto (2020) No. ISSN 2548-8368	Berdasarkan pengujian, dapat disimpulkan bahwa algoritma <i>FP-Growth</i> menjadi solusi untuk mengatasi kekurangan pada apriori yaitu gangguan batasan area dalam

	Produk Internet		<p>penelusuran aturan asosiasi pada jangkauan semua area yang telah terpasang produk Oxygen. Dalam penelitian ini juga diketahui bahwa nilai ID Produk dan nilai ID Area menyesuaikan perhitungan algoritma <i>FP-Growth</i>, sebab makin besar nilai <i>minimum support</i> yang di <i>input</i> maka jumlah aturan asosiasi yang dimunculkan semakin banyak.</p>
6	Rekomendasi Pemilihan Peminjaman Buku Favorit Favorit di	Hasna Rasyidah, Saruni Dwiasnati (2021) No. ISSN 2503-	<p>Pada dataset peminjaman buku yang diterapkan nilai <i>minimum support</i></p>

	Perpustakaan Menggunakan Algoritma <i>FP-Growth</i>	054X	20% dan <i>minimum confidence</i> 70% maka didapat 22 aturan bisa ditentukan jikalau meminjam buku yang genre Psikologi dan Filsafat maka kemungkinan besar juga akan meminjam buku genre Agama.
7	Penerapan Data Mining Algoritma <i>FP-Growth</i> Untuk Persediaan <i>Sparepart</i> Pada Bengkel Motor (Studi Kasus Bengkel Sinar Servis)	Boby Septia Pranata, Dito Putro Utomo (2020) No. ISSN 2722-0524	Dapat meningkatkan persediaan <i>sparepart</i> pada bengkel motor berdasarkan data. Hasil pengujian memunculkan bahwa <i>sparepart x.y</i> dengan nilai 8.9 tertinggi dan terendah atas proses <i>FP-Growth</i> data bengkel sinar servis.
8	<i>Association Rules Analysis on FP-</i>	Supiyandi, Mochammad Iswan	<i>Data mining</i> dapat diimplementasikan

	<i>Growth Method in Predicting Sales</i>	Perangin-angin, Andre Hasudungan Lubis, Ali Ikhwan, Mesran, Andysah Putera Utama Siahaan (2017) No. ISSN 2455-1457	menggunakan transaksi penjualan untuk menemukan pola tren kombinasi antar item-item. Dapat menentukan stok penyediaan barang dan teknik penjualan yang efisien dan tepat.
9	<i>Market Basket Analysis Using Apriori and FP-Growth for Analysis Consumer Expenditure Patterns at Berkah Mart in Pekanbaru Riau</i>	Mustakim, Della Maulina Herianda, Ahmad Ilham, Achmad Daeng GS, Folkes E. Laumal, Nuning Kurniasih, Akbar Iskandar, Gloria Manulangga, Ida Bagus Ary Indra Iswara and Robbi Rahim (2018) No.DOI :10.1088/1742-6596/1114/1/01213	Hasil menunjukkan bahwa pengaplikasian <i>FP-Growth</i> dan <i>Apriori algorithm</i> untuk analisis pola belanja konsumen dapat meningkatkan keuntungan namun sangat disarankan menggunakan <i>FP-Growth algorithm</i> yang untuk proses pencarian aturan

		1	secara maksimal dan efisien.
10	<i>Market Basket Analysis using FP-Growth Association Rule on Textile Industry</i>	Kadek Darmaastawan, Komang Oka Saputra, and Ni Made Ary Esta Dewi Wirastuti.	<i>Data mining</i> dibutuhkan dalam penelitian ini menggunakan <i>FP-Growth Association Rule</i> dari transaksi data penjualan <i>online</i> tekstil di Bali. <i>FP-Growth association rule</i> terdiri atas beberapa tahap, <i>frequent item (support items) discovery</i> , <i>FP-Tree construction</i> , <i>frequent itemset (support itemset) discovery</i> , <i>association rules discovery</i> , dan <i>lift ratio calculation</i> .
11	<i>Market Basket</i>	Febrian Teguh	Aplikasi algoritma

	<p><i>Analysis</i></p> <p>Menggunakan</p> <p>Algoritma <i>FP-Growth</i></p> <p>(Kasus Data Peminjaman Buku di Perpustakaan Universitas Islam Bandung)</p>	<p>Raharjo, Teti Sofia Yanti, Abdul Kudus (2017)</p> <p>No. ISSN 2460-6456</p>	<p><i>FP-Growth</i> dengan <i>minimum support</i> 3% dari keseluruhan transaksi ditemukan 14 kode buku yang sering dipinjam dan 10 <i>itemssets</i> yang sering dipinjam secara bersamaan dari data peminjaman buku di Perpustakaan Universitas Islam Bandung.</p>
12	<p>Algoritma <i>FP-Growth</i> untuk</p> <p>Menganalisa</p> <p>Frekuensi</p> <p>Pembelian Gas Elpiji 3 Kg</p>	<p>Erlin Elisa, Nurul Azwanti</p> <p>No. ISSN 2549-6824</p>	<p>Hasil yang diperoleh berdasarkan hasil penghitungan algoritma <i>FP-Growth</i> dan sistem yang dibuat dapat dilakukan pengaturan penyediaan stok dan distribusi kepada</p>

			<p>konsumen.</p> <p>Kemudian pengimplementasian <i>Algoritma FP-Growth</i> pada teknik <i>data mining</i> lebih efisien dan mempermudah proses penyusunan kecenderungan pembelian pada pangkalan elpiji UD. Maju Bersama dimana yang terjual paling banyak atau terbeli pada minggu 1 dan 2 pada setiap bulannya dengan nilai tertinggi <i>support</i> 66,67% <i>confidence</i> 100%</p>
13	Implementasi <i>Algoritma FP-Growth</i> Pada	Nurasiah (2021) No. ISSN 2722-7987	<i>Algoritma FP-Growth</i> dapat diterapkan untuk

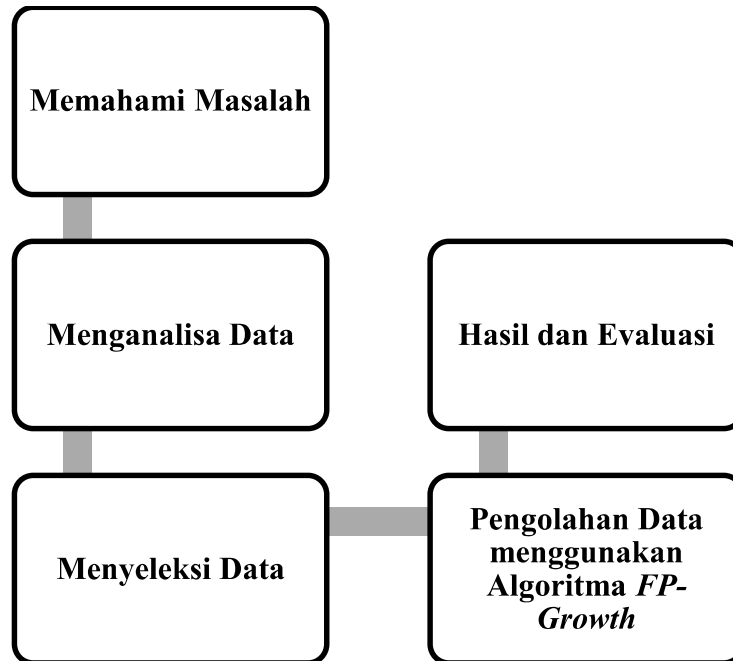


	Pengenalan Pola Penjualan		memperoleh pola penjualan ataupun korelasi produk pada CV. Bagus Alam Sejahtera. Hasil percobaan algoritma <i>FP-Growth</i> pada <i>rapidminer</i> memanifestasikan kalau penjualan BS ( <i>Blonde</i> Sedang) dan BA (Bika Ambon) dengan nilai <i>confidence</i> 50% dan nilai <i>support</i> 17% yang bersumber dari data tahun 2016-2018 di CV. Bagus Alam Sejahtera.
14	Teknik <i>Data Mining</i> Untuk Penentuan Paket Hemat Sembako Dan Kebutuhan	Amelia Nastuti (2019) No. ISSN 2615-1855	Proses pelatihan dengan 15 data sampel telah dianalisa melalui tahapan

	Harian Dengan Menggunakan Algoritma <i>FP-Growth</i> (Studi Kasus Di Ulfamart Lubuk Alung)		<i>preprocessing</i> untuk menemukan kombinasi item produk yang dapat dijadikan sebuah paket hemat, dimana memperoleh 5 rekomendasi paket hemat.
15	Implementasi <i>Data Mining</i> Pada Penjualan Produk Di CV. Cahaya Setya Menggunakan Algoritma <i>FP-Growth</i>	Wahyu Nur Setyo, Sukma Wardhana (2019) No. ISSN 2655-5018	Bahwa semakin kecil nilai <i>minimum support</i> maka bertambah banyak aturan asosiasi yang diperoleh, dan sebaliknya .

### 2.2.7 Kerangka Pemikiran

Dibawah ini terdapat kerangka penelitian bentuk diagram yang merincikan tahapan – tahapan sebagai berikut :



**Gambar 2.2** Kerangka Pemikiran

Keterangan kerangka pemikiran diatas adalah :

Masalah yang terjadi pada Kiki Ms. Glow adalah data transaksi penjualan pada Kiki Ms. Glow belum diolah sehingga terjadi penimbunan data selama penjualan, maka dari itu peneliti mengambil data dan menganalisa serta menyeleksi data yang akan dilakukan penggalan data menggunakan teknik algoritma *FP-Growth*.

Algoritma yang dipakai dalam penelitian ini menggunakan *FP-Growth*. *FP-Growth* adalah salah satu algoritma yang tergolong *assosication rule*. Algoritma *FP-Growth* dibagi menjadi tiga tahap utama, yaitu:

1. Tahap Pembangkitan *Conditional Pattern Base*.

*Conditional Pattern Base* yaitu sub basis data yang berisi lintasan prefix (*prefix path*) dan pola akhiran (*suffix pattern*). Pencetusan *conditional pattern base* diperoleh melalui *FP-Tree* yang telah dibuat sebelumnya.

2. Tahap Pembangkitan *Conditional FP-Tree*.

Di tahap ini *support count* dari tiap item ke *conditional pattern base* ditambahkan, kemudian tiap item yang mempunyai total *support count* lebih besar atau sama dengan *minimum support count* akan dimunculkan dengan *conditional FP-Tree*.

3. Tahap Pencarian *Frequent Itemset*

Bilamana hasil dari *Conditional FP-Tree* merupakan lintasan tunggal (*single path*), maka diperoleh *frequent itemset* dengan mengkombinasi item untuk setiap *conditional FP-Tree*. Jika bukan lintasan tunggal, maka dilanjutkan pembangkitan *FP-Growth* secara rekursif.

1. Hasil dan Evaluasi

Diharapkan bisa menyampaikan hasil serta evaluasi bagi pegiat usaha *retail* untuk menaikkan penjualan serta promosi produk-produk kosmetik.

### 2.2.7 Hipotesis

Hipotesis penelitian berdasarkan kerangka penelitian adalah sebagai berikut ini:

1. Analisis *association rule* menggunakan algoritma *FP-Growth* dapat menemukan *FP-Tree* dari nilai *support* dan *confidence* dari transaksi penjualan selama ini.
2. Penggalan data diduga dapat mempermudah pebisnis dalam persediaan barang dan menentukan tingkat pembelian konsumen.