

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Tinjauan Penelitian Terdahulu**

Penelitian dibidang ini telah banyak dilakukan, pembahasan topik yang diangkat untuk dijadikan dasar serta sumber data. Hal ini bertujuan untuk mendukung penelitian yang sedang dilakukan. Selain itu untuk mengajukan gagasan atau argumen diperlukan data yang relevan dengan penelitian sehingga dapat dikatakan valid. Tujuan lain dilakukannya studi terhadap penelitian terdahulu adalah untuk menggali kekurangan dan kelebihan suatu metode yang akan digunakan. Dengan begitu peneliti dapat menghindari atau mengatasi permasalahan terkait.

Oleh sebab itu sebelum memulai penelitian perlu dilakukan eksplorasi dan studi pustaka baik melalui internet maupun jurnal – jurnal yang relevan dengan topik penggalan data opini atau Analisis Sentimen. Berikut adalah beberapa penelitian yang dijadikan tinjauan pustaka. Pada beberapa halaman selanjutnya disediakan tabel perbandingan antar-penelitian yang dapat dilihat pada Tabel 2.1

##### **2.1.1. Implementasi Data Mining Dengan Algoritma Apriori (Studi Kasus: Bengkel Akmal Jaya Motor Purwodadi)**

Penelitian pertama oleh Ristianingrum dan Sulastri (2017) dengan judul “Implementasi Data Mining Dengan Algoritma Apriori (Studi kasus: Bengkel Akmal Jaya Motor Purwodadi). Penelitian ini menerapkan teknik Data Mining

pada Bengkel Akmal Jaya yang menghasilkan sebuah aturan asosiatif yang nantinya berpengaruh pada analisis hasilnya, sehingga pemilik bengkel mengetahui seberapa tinggi frekuensi suku cadang dan jasa apa yang paling sering terjadi. Maka dari itu dilakukan analisis dan pengujian yang bisa memberikan informasi mengenai pola transaksi dan mengetahui jasa dan suku cadang apa yang sering muncul dari bengkel Akmal Jaya Motor sehingga bisa membantu pemilik dalam mengambil keputusan untuk melakukan pembelian suku cadang apa saja yang perlu dibeli.

### **2.1.2 Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis Web**

Penelitian ke-2 oleh Agus Nursikuwagus (2016) dengan judul penelitian “Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis Web”. Penelitian ini bertujuan untuk melaporkan hasil dari implementasi algoritma apriori untuk analisa penjualan. Penelitian ini didukung dengan metodologi pengembangan perangkat lunak yang disebut *linear incremental development* Pada proses yang dilakukan pada implementasi perangkat lunak ini terdapat beberapa langkah seperti penentuan masalah, identifikasi dan penyelesaian resiko, pengembangan dan test, serta perencanaan siklus berikutnya. Setiap langkah pada tahapan tersebut digunakan untuk memetakan permasalahan yang ada hingga terbentuknya suatu aplikasi sesuai kebutuhan. Sebagai hasil dari penelitian ini adalah aplikasi berdasarkan algoritma apriori yang terdiri dari pemilihan frekuensi Itemset dan aturan asosiasi. Keluaran dari perangkat lunak ini digunakan untuk data acuan stok barang serta mengotomatisasi perhitungan

analisis penjualan barang. Sedangkan *tools* yang digunakan untuk merancang sistem ini adalah UML (*Unified Modelling Language*) dan menggunakan beberapa diagram dari UML tersebut, seperti: *Use case Diagram*, *Class Diagram*, dan *Sequential Diagram*.

### **2.1.3 Implementasi Algoritma Apriori Aturan Keterkaitan Data Untuk Analisis Keranjang Belanja Sistem Persediaan Obat Pada Apotek Perdos Farma Makassar**

Penelitian ke-3 oleh Ulvah dengan judul “Implementasi Algoritma Apriori Aturan Keterkaitan Data Untuk Analisa Keranjang Belanja Sistem Persediaan Obat Pada Apotek Perdos Farma Makassar”. Penelitian ini menganalisa Populasi yang merupakan objek atau subjek memiliki nilai dan ciri-ciri tertentu yang sifatnya nyata baik berupa benda hidup, benda mati dan manusia yang dapat menghasilkan informasi bagi peneliti untuk kemudian digunakan dalam penelitian seperti halnya jumlah obat dan alat kesehatan yang ada pada apotek perdos farma makassar yang penulis temukan sebanyak 1406 item. Sedangkan sampel merupakan beberapa bagian dari populasi yang diambil untuk kepentingan penelitian dengan menggunakan teknik *random sampling* Karena jumlah populasi sebanyak 1406. Penelitian menghasilkan 6 *rules* terbaik dengan nilai *minimum support* sebesar 15%, dan nilai *confidence* tertinggi dari 6 *rules* terbaik sebesar 72% dengan tingkat keakuratan nilai *lift* tertinggi 6,27%. Dari hal tersebut dapat diketahui bahwa hanya 14 jenis obat yang paling banyak terjual berdasarkan 100 transaksi penjualan dari 312 jenis obat pada Apotek Perdos Farma Makassar.

**Tabel 2.1.** Penelitian Terdahulu

No.	Peneliti	Judul Penelitian	Perumusan Masalah	Manfaat
1.	Ristianingrum dan Sulastri (2017)	Implementasi Data Mining Menggunakan Algoritma Apriori	Bagaimana cara menerapkan data mining menggunakan algoritma apriori untuk melakukan pembelian suku cadang apa saja yang perlu dibeli oleh bengkel Akmal Jaya Motor	Membantu AHASS Akmal Jaya Motor dalam menentukan produk suku cadang (spare part) apa saja yang harus dibeli oleh AHASS Akmal Jaya Motor.
2.	Agus Nursikuwagus (2016)	Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis Web	Bagaimana cara mengimplementasikan algoritma apriori untuk analisa penjualan berbasis web	Dengan hasil yang diperoleh dapat digunakan untuk membantu pengambil keputusan. implementasi ini bisa membantu untuk mengadakan stok produk yang banyak disukai oleh pembeli, dan menambah persediaan produk.

Lanjutan 2.2 Penelitian Terdahulu (Lanjutan)

---

3.	Ulvah (2018)	Implementasi Algoritma Apriori Aturan Keterkaitan Data Untuk Analisa Keranjang Belanja Sistem Persediaan Obat Pada Apotek Perdos Farma Makassar	Bagaimana cara mengimplementasikan algoritma apriori untuk persediaan obat pada apotek perdos farma Makassar	Dengan didapatnya rules terbaik maka pihak apotek dapat menggunakan rules tersebut dalam membuat strategi strategi untuk meningkatkan pesediaan agar dapat mengurangi kerugian biaya dan dapat memenuhi kebutuhan konsumen.
----	--------------	---	--	---

Tabel.2.1 Penelitian Terdahulu

## 2.2 Teori Dasar Yang Digunakan

Pada penelitian ini terdapat beberapa teori dasar yang digunakan sebagai acuan dan referensi terkait pembahasan mengenai pengolahan data transaksi penjualan produk pada kafe. Sumber di dapat dari jurnal yang relevan dengan penelitian ini serta dari studi pustaka melalui buku dan internet.

### 2.2.1 Kafe Kopiganes

Kopiganes merupakan salah satu kafe di Surabaya, kopiganes adalah rumah kopi dan galeri seni berkonsep budaya jawa yang berdiri pada tanggal 21 maret 2015. Kopiganes bernuansa seperti di pedesaan yang jauh dari hingar bingar kota Surabaya, berlokasi di bagian timur kota Surabaya, tepatnya di jalan Keputih Tegal Timur 2 nomor 116 Sukolilo. Kopiganes merupakan usaha yang

memanfaatkan peluang berkembangnya gaya hidup masyarakat dari kegemaran mengkosumsi kopi. Dengan target pemasaran dikalangan semua umur tetapi lebih ditekankan kepada remaja dikarenakan lebih banyak pengkonsumsi di kafe kopiganes ini remaja yang suka nongkrong di tempat tersebut. Pada sistem ini, pengambilan data dilakukan dengan datang langsung ke kafe, dan melakukan wawancara serta pengambilan data transaksi penjualan produk.

### **2.2.2 Strategi Penjualan**

Strategi penjualan adalah upaya menjual suatu produk, baik itu barang atau jasa, dengan menggunakan pola rencana dan taktik tertentu sehingga jumlah penjualan akan semakin meningkat. Pengertian strategi penjualan juga dapat diartikan sebagai rangkaian atau upaya yang dilakukan oleh perusahaan untuk mencapai suatu target atau tujuan tertentu, karena potensi untuk menjual proposisi terbatas pada jumlah orang yang mengetahui hal tersebut. Strategi penjualan punya peranan penting dalam sebuah perusahaan atau bisnis karena berfungsi untuk menentukan nilai ekonomi perusahaan, baik itu suatu harga barang maupun jasa. Ada tiga faktor yang mempengaruhi penentu nilai harga barang dan jasa yaitu produksi, pemasaran, dan konsumsi. Dalam hal ini strategi penjualan menjadi bagian penting yang dapat menghubungkan antara kegiatan produksi dan konsumsi.

Ada beberapa pengertian strategi penjualan. Dibawah ini pengertian pemasaran diberikan oleh beberapa ahli sebagai berikut:

- Menurut Kotler dan Amstrong (2008)

Pengertian strategi penjualan adalah logika pemasaran dimana unit bisnis berharap untuk menciptakan nilai dan memperoleh keuntungan dari hubungannya dengan konsumen.

- Menurut Kurtz (2008)

Pengertian strategi penjualan adalah keseluruhan program karakteristik perusahaan dalam menentukan target pasar yang dapat memuaskan konsumen dengan membangun kombinasi elemen atau item dari *marketing mix* (produk, distribusi, promosi, dan harga)

- Menurut Philip Kotler (1997)

Pengertian strategi penjualan adalah suatu mindset pemasaran yang akan digunakan untuk mencapai tujuan pemasaran, dimana di dalamnya terdapat strategi rinci mengenai pasar sasaran, penetapan posisi, bauran pemasaran, dan budget untuk pemasaran.

- Menurut Tjiptono

pengertian strategi penjualan adalah alat fundamental atau pokok yang dirancang untuk mencapai tujuan perusahaan dengan mengembangkan keunggulan daya saing yang berkesinambungan melewati pasar yang dimasuki, dan program strategi penjualan yang digunakan untuk melayani pasar target tersebut.

- Menurut Stanton

Pengertian strategi penjualan adalah sesuatu yang melingkupi semua sistem yang memiliki hubungan dengan tujuan untuk merencanakan dan menentukan harga hingga mempromosikan dan menyalurkan produk (barang atau jasa) yang dapat memuaskan konsumen.

Fungsi dari strategi penjualan:

1. Meningkatkan motivasi untuk melihat masa depan
2. Koordinasi pemasaran yang lebih efektif
3. Merumuskan tujuan dari perusahaan
4. Pengawasan kegiatan pemasaran

Tujuan dari strategi penjualan:

1. Untuk meningkatkan kualitas koordinasi antar individu dalam tim pemasaran
2. Sebagai alat ukur hasil pemasaran berdasarkan standar prestasi yang telah ditentukan
3. Sebagai dasar logis dalam mengambil keputusan pemasaran
4. Untuk meningkatkan kemampuan dalam beradaptasi bila terjadi perubahan dalam pemasaran

## Konsep strategi penjualan

### 1. Segmentasi Pasar

Setiap konsumen pasti memiliki kebutuhan dan kebiasaan yang berbeda. Perusahaan harus melakukan klasifikasi pasar yang sifatnya heterogen menjadi satuan-satuan pasar yang bersifat homogen.

### 2. *Market Positioning*

Tidak ada perusahaan yang bisa menguasai seluruh pasar. Itulah alasannya mengapa perusahaan harus mempunyai pola karakteristik untuk mendapatkan posisi kuat dalam pasar, yaitu memilih segmen yang paling menguntungkan dan mengerti akan kondisi pasar saat ini.

### 3. *Market Entry Strategy*

Ini adalah strategi perusahaan untuk bisa masuk pada segmen pasar tertentu. Beberapa cara yang sering dilakukan adalah membeli perusahaan, internal development, dan kerja sama dengan perusahaan lain.

### 4. *Marketing Mix Strategy*

kumpulan dari beberapa variabel yang telah digunakan perusahaan untuk mempengaruhi tanggapan konsumen. Beberapa variabel tersebut diantaranya produk, harga, tempat, promo, peserta, proses, bukti fisik seseorang

## 5. *Timing Strategy*

Pemilihan waktu dalam melakukan penjualan juga sangat penting untuk diperhatikan. Perusahaan perlu melakukan berbagai persiapan yang baik di bidang produksi yang harus sudah terpenuhi, dan menentukan waktu yang tepat untuk mendistribusikan produk ke pasar.

### 2.2.3 Data Mining

Data Mining ialah serangkaian proses operasi yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi dari data yang bermanfaat dan pengetahuan yang terkait dari berbagai database atau basis data (Turban dkk. 2005). Terdapat beberapa 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*) dan *data archaeology* dan *data dredging* (Larose,2005)

Kelebihan dari data mining ialah untuk mencari informasi yang berharga dari database, seperti contoh penambangan emas atau logam mulia dari sumber lahannya, teknologi ini dipakai untuk :

1. Prediksi trend dan sifat-sifat bisnis, dimana data mining mengotomatisasi proses pencarian informasi prediksi di dalam database atau data yang sangat banyak.

2. Penemuan pola yang belum diketahui sebelumnya, dimana data mining memproses data yang sangat besar, kemudian mengidentifikasi pola tersebut dalam satu proses.
3. Data mining berguna untuk membuat keputusan yang akurat, terutama dalam strategi kedepannya.

Berikut ini adalah beberapa definisi data mining dari beberapa sumber (Larose,2005):

- Data mining adalah proses menemukan sesuatu yang bermakna dari suatu korelasi, pola dan tren yang ada dengan cara memilah-milah data berukuran besar yang disimpan dalam memori, menggunakan teknologi pengenalan pola serta teknik matematika dan statistik.
- Data mining adalah analisis pengamatan database untuk menemukan hubungan yang tidak terduga dan untuk meringkas data dengan cara atau metode baru yang dapat dimengerti dan bermanfaat kepada pemilik data.
- Data mining adalah serangkaian bidang ilmu studi disiplin tinggi yang menggabungkan teknik pembelajaran dari mesin (*machine learning*), pengenalan pola (*pattern recognition*), statistik, basis data yang besar, dan visualisasi untuk mengatasi masalah ekstraksi informasi dari database.
- Data mining diartikan sebagai suatu proses ekstraksi informasi berguna dan potensial dari sekumpulan data yang terkandung didalamnya suatu basis data.

## **Fungsi Data Mining**

Data mining mempunyai fungsi yang sangat penting untuk membantu mendapatkan informasi yang berfungsi untuk meningkatkan pengetahuan bagi para pengguna. Pada dasarnya, data mining mempunyai empat fungsi dasar yang harus diketahui yaitu:

1. Fungsi Prediksi (*prediction*). Proses untuk menemukan pola dari data dengan menggunakan beberapa variabel untuk memprediksikan variabel lain yang tidak diketahui jenis atau nilainya.
2. Fungsi Deskripsi (*description*). Proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data.
3. Fungsi Klasifikasi (*classification*). Klasifikasi merupakan suatu proses untuk menemukan pola atau fungsi untuk menggambarkan konsep dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang sangat berpengaruh serta dapat memprediksikan kecenderungan data tersebut pada masa yang akan datang.
4. Fungsi Asosiasi (*association*). Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data.

## **Proses Data Mining**

Proses yang pada umumnya dilakukan oleh data mining antara lain: deskripsi, prediksi, estimasi, klasifikasi, clustering dan asosiasi. Proses data mining dijelaskan sebagai berikut (Larose, 2005):

### *a. Deskripsi*

Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Dalam proses data mining aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (*knowledge*) pada sistem. Tugas deskriptif merupakan tugas data mining yang sering dibutuhkan pada teknik *postprocessing* untuk melakukan validasi dan menjelaskan hasil dari proses data mining. *Postprocessing* merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.

### *b. Prediksi*

Prediksi hampir memiliki kesamaan dengan klasifikasi, akan tetapi data yang diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga dalam beberapa bulan yang akan datang.

### *c. Estimasi*

Estimasi hampir sama dengan prediksi, kecuali variabel target estimasi lebih ke arah *numerik* dari pada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

Selanjutnya, pada peninjauan berikutnya estimasi mempunyai nilai dari variabel yang telah ditargetkan lalu dibentuk berdasarkan nilai variabel prediksi. Sebagai contoh, dilakukan estimasi transaksi penjualan pada menu kafe berdasarkan nama menu, kode menu, harga menu, dan jumlah penjualan. Hubungan antara transaksi penjualan dan nilai variabel prediksi dalam proses pembelajaran yang akan menghasilkan pola berdasarkan standart dari aturan estimasi.

#### *d. Klasifikasi*

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

#### *e. Clustering*

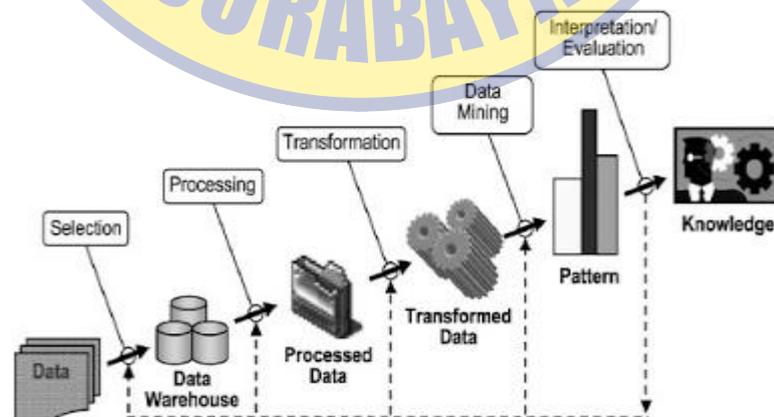
*Clustering* ialah penggabungan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah kluster adalah kumpulan *record* data yang memiliki kesamaan suatu dengan yang lainnya dan memiliki ketidak samaan dengan *record* data dalam kluster lain. Tujuannya adalah untuk menghasilkan penggabungan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kesamaan pada objek dalam suatu cluster dan semakin besar pula perbedaan tiap cluster maka kualitas analisis cluster semakin baik.

f. Asosiasi

Tugas asosiasi dalam data mining ialah menemukan atribut atau item yang muncul dalam satu waktu yang sama. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (*market basket analysis*). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur seberapa tinggi frekuensi hubungan antara dua atau lebih atribut atau item.

**Tahapan Data Mining**

Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap *preprocessing* untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan *output* berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik. Secara detail dijelaskan sebagai berikut (Fayyad, 1996):



Gambar.2. 1 Tahapan Data Mining

### 1. *Data selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

### 2. *Pre-processing / cleaning*

Sebelum proses data mining dapat dikerjakan, perlu dilakukannya proses *cleaning* pada data yang menjadi fokus data mining. Proses *cleaning* mencakup antara lain membuang data yang mirip, memeriksa data yang inkonsisten, dan memperbaiki kesalahan yang terdapat pada data.

### 3. *Transformation*

*Coding* adalah proses modifikasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam KDD merupakan proses yang sangat bergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

### 4. *Data mining*

*Data mining* adalah proses mencari pola, model atau informasi penting dalam data yang telah dipilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat beragam dan

mempunyai banyak variasi. Pemilihan teknik, metode atau algoritma yang tepat sangat berpengaruh pada tujuan dan proses KDD secara keseluruhan.

#### 5. *Interpretation / evaluation*

Pola informasi yang dihasilkan dari proses data mining harus digambarkan dalam bentuk yang sangat mudah dimengerti oleh pengguna atau pihak yang berkepentingan. Tahapan ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahapan ini mencakup evaluasi apakah pola atau informasi yang ditemukan berbanding terbalik dengan fakta atau hipotesis yang ada sebelumnya.

#### 2.2.4 *Association Rule* (Aturan Asosiasi)

*Association rule mining* adalah suatu prosedur untuk mencari hubungan antar item dalam suatu data set yang ditentukan (Han dan Kamber, 2006). *Association rule* meliputi dua tahapan yaitu mencari kombinasi yang paling sering muncul dari suatu itemset dan mendefinisikan *condition* dan *result* (untuk conditional association rule). Masalah mendasar dari analisis asosiasi adalah bagaimana menemukan kaidah dalam bentuk set1 ke set 2 (Possas, Wagner, Marcio dan Rodolfo, 2000).

*Association Rule* atau aturan asosiasi adalah teknik data mining untuk menemukan aturan asosiatif atau pola kombinasi dari suatu item. Dalam aturan asosiasi yang berbentuk “*if...then...*” atau “jika...maka...” dengan kata lain aturan asosiasi berkenaan atau analisis afinitas dengan studi tentang “jika maka”.

Karena awalnya berasal dari studi tentang database transaksi pelanggan untuk menentukan kebiasaan suatu produk apa, maka aturan asosiasi juga sering dinamakan *market basket analysis* atau aturan yang menyatakan asosiasi antara beberapa atribut (Han, 2006).

*Association rule* merupakan bentuk jika “kejadian sebelumnya” kemudian “konsekuensinya” (*if antecedent, then consequent*), yang diikuti dengan perhitungan aturan *support* dan *confidence*. Contoh dalam sebuah transaksi pembelian suatu produk di sebuah kafe didapat *association rule* kopi → roti bakar. Pelanggan yang membeli kopi berkemungkinan besar pelanggan tersebut juga akan membeli roti bakar dalam waktu yang bersamaan.

Pada proses *metode association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang dapat diperoleh dari hasil pengolahan data dengan proses perhitungan tertentu yaitu:

- *Support*: suatu frekuensi yang menentukan seberapa besar tingkat dominasi suatu item/itemset dari keseluruhan transaksi. Pada frekuensi ini berguna untuk menentukan apakah suatu itemset tersebut berpengaruh pada *confidence*-nya (misalnya suatu atribut dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi yang menunjukkan bahwa item A dan B dibeli secara bersamaan)
- *Confidence*: suatu ukuran yang menunjukkan hubungan antara 2 item secara *conditional* (misal, seberapa sering item B jika orang

membeli item A). Misalkan item A dan B dalam total transaksi, maka *support* (A) adalah jumlah transaksi yang ada A dibagi total transaksi yang disebut *support* 1-item set, dan *support* (A dan B) adalah transaksi yang ada A dan B dibagi total transaksi yang disebut *support* 2-item set. Demikian selanjutnya untuk item barang yang lebih banyak. Sedangkan *confidence* dianalisa mulai dari 2 barang karena berhubungan dengan keinginan membeli barang secara bersamaan. Rumusnya adalah *confidence* (A → B) sama dengan *support* (A dan B) dibagi *support* (A). Perhatikan bahwa akan beda perhitungan dengan *confidence* (B → A). Demikian untuk item barang yang lebih banyak. Definisi formal untuk kedua ukuran ini dapat dirumuskan:

$$\text{Support: } (X \rightarrow Y) = \sigma(X \cup Y) / N$$

$$\text{Confidence: } (X \rightarrow Y) = \sigma(X \cup Y) / \sigma(X)$$

Kedua frekuensi ini berfungsi untuk menentukan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua parameter tersebut. Bila suatu pola memenuhi kedua parameter kedua minimum parameter yang sudah ditentukan, maka pola tersebut dapat disebut sebagai *interesting rule*. Kedua frekuensi ini nantinya berfungsi untuk menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh user. Batasan tersebut umumnya terdiri dari *min\_support* dan *min\_confidence*. Contoh (suatu

*association rule*), *If A then B* [*support*=30%, *confidence*=70%], dimana A dan B adalah kumpulan item yang dibeli oleh konsumen perusahaan Z. Artinya: item A dan B dibeli bersamaan sebesar 30% dari keseluruhan data transaksi yang dianalisis dan 70% dari semua konsumen yang membeli item A juga membeli item B. Dari contoh di atas, jika *support*-nya  $\geq \text{min\_support}$  dan *confidence*-nya  $\geq \text{min\_confidence}$ , maka *rule* tersebut bisa dikatakan sebagai *interesting rule*.

### 2.2.5 Algoritma Apriori

Algoritma apriori termasuk jenis aturan pada data mining. Menurut Dan Toomey (2014:49-50), Apriori adalah algoritma kelas yang membantu mempelajari aturan asosiasi. Ini bekerja melawan transaksi. Algoritma mencoba untuk menemukan himpunan bagian yang umum dalam kumpulan data. Ambang batas minimum harus dipenuhi agar asosiasi dapat dikonfirmasi.

Analisis asosiasi adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Algoritma yang paling populer dikenal dengan nama Apriori karena paradigma *generate and test*, yaitu menentukan kandidat kombinasi itemset tersebut memenuhi syarat *support minimum* (Santosa,2007).

Kombinasi item yang memenuhi syarat tersebut disebut *frequent itemset*, yang nantinya dipakai untuk membuat aturan-aturan yang memenuhi syarat

confidence minimum. Suatu aturan assosiatif dapat diketahui dengan dua parameter, yaitu *support* dan *confidence*. *Support* (nilai penunjang) adalah frekuensi tertinggi kombinasi item tersebut dalam basis data, sedangkan *confidence* (nilai kepastian) adalah kuatnya frekuensi hubungan antar item dalam aturan assosiasi (Lutfi, dkk, 2009).

Proses utama yang dikerjakan dalam algoritma apriori adalah untuk mendapat *frequent itemset* yaitu:

- *Join* (penggabungan)

Proses ini dilakukan dengan cara pengkombinasian item dengan item lainnya sehingga tidak bisa terbentuk kombinasi lagi

- *Prune* (pemangkasan)

Proses pemangkasan yaitu hasil dari item yang telah dikombinasikan kemudian dipangkas dengan menggunakan *minimum support* yang telah ditentukan.

Pada algoritma Apriori ditentukananya kandidat apa saja yang mungkin muncul dengan cara menentukan *minimum support* dan *minimum confidence*. *Support* adalah nilai penunjang atau persentase frekuensi tertinggi kombinasi sebuah item dalam basis data.

Rumus *support* adalah sebagai berikut:

$$\text{Support (A,B)} = P (A \cap B)$$

$$\text{Support (A,B)} =$$

$$\frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{total transaksi}} \times 100\%$$

Sedangkan *confidence* adalah nilai kepercayaan yaitu kuatnya hubungan antar item dalam sebuah Apriori. *Confidence* dapat dicari setelah pola frekuensi munculnya sebuah item ditemukan. Berikut rumus *confidence*:

Confidence = P(B|A) =

$$\frac{\sum \text{transaksi mengandung } A \text{ dan } B}{\sum \text{total transaksi A}} \times 100\%$$

Cara kerja apriori :

1) Tentukan nilai *minimum support* perhitungan itemset 1 : hitung itemset dari *support* (transaksi yang memuat seluruh item atau atribut) dengan men-*scan* database untuk 1-itemset, setelah 1-itemset didapatkan, dari 1-itemset apakah memenuhi standart nilai *minimum support* yang telah disetting, apabila telah memenuhi *minimum support*, 1-itemset tersebut akan menjadi pola *frequent* tinggi dan dilanjutkan kelangkah berikutnya. Iterasi 2 : untuk mendapatkan 2-itemset, dilakukanya kombinasi dari *k-itemset* dari iterasi item 1, kemudian *scan* database lagi untuk hitung item-item yang memuat *minimum support*. itemset yang memenuhi standart nilai *minimum support* akan dipilih sebagai pola *frequent* tinggi dari kandidat itemset.

2) Tetapkan nilai *k-itemset* dari *support* yang telah memenuhi *minimum support* dari *k-itemset*. melakukan proses untuk proses perhitungan selanjutnya hingga tidak ada lagi *k-itemset* yang memenuhi *minimum support*. Algoritma

Apriori sendiri termasuk ke dalam data mining asosiasi atau *association rule mining* yang teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Kelebihan dari algoritma apriori ini adalah lebih sederhana dan dapat menangani data yang besar. tentunya berpengaruh terhadap banyaknya item yang diproses serta mudah di pahami struktur kerja dan implementasinya. Apriori memiliki kelemahan yaitu harus melakukan scan database setiap kali iterasi sehingga terlalu memenuhi kapasitas memori, sehingga waktu yang diperlukan bertambah dengan makin banyak iterasi.

