

**PENERAPAN METODE ANALISIS ASOSIASI PADA DATA
PENJUALAN BARANG MENGGUNAKAN ALGORITMA APRIORI
(Studi Kasus pada Supermarket Robinson Plaza Andalas)**

SKRIPSI

*Diajukan untuk memenuhi Sebagian Dari Syarat Memperoleh Gelar Sarjana
Program Studi Pendidikan Teknik Informatika*



**Oleh
MARISA RAHMA DANTI
NIM: 55736.2010**

**PROGRAM STUDI PENDIDIKAN TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRONIKA
FAKULTAS TEKNIK
UNIVERSITAS NEGERI PADANG
2014**

PERSETUJUAN PEMBIMBING

SKRIPSI

Judul : Penerapan Metode Analisis Asosiasi pada Data Penjualan Barang
Menggunakan Algoritma Apriori (Studi Kasus pada Supermarket
Robinson Plaza Andalas)

Nama : Marisa Rahma Danti

NIM : 55736

Program Studi : Pendidikan Teknik Informatika

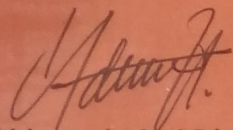
Jurusan : Teknik Elektronika

Fakultas : Teknik

Padang, September 2014

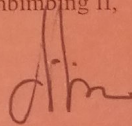
Disetujui oleh :

Pembimbing I,



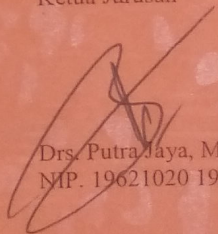
Muhammad Adri, S.Pd, MT
NIP. 19750514 200003 1 001

Pembimbing II,



Oktoria, S.Pd, MT
NIP. 19831010 200801 1 017

Ketua Jurusan



Drs. Putra Jaya, MT
NIP. 19621020 198602 1 001

PENGESAHAN TIM PENGUJI

Nama : Marisa Rahma Danti
NIM : 55736

Dinyatakan lulus setelah mempertahankan skripsi di depan Tim Penguji
Program Studi Pendidikan Teknik Informatika
Jurusan Teknik Elektronika
Fakultas Teknik
Universitas Negeri Padang
dengan judul

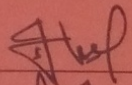
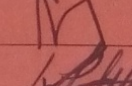
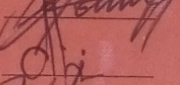
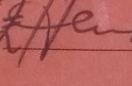
**Penerapan Metode Analisis Asosiasi pada Data Penjualan Barang
Menggunakan Algoritma Apriori
(Studi Kasus pada Supermarket Robinson Plaza Andalas)**

Padang, September 2014

Tim Penguji

Tanda Tangan

1. Ketua : Asrul Huda, S.Kom, M.Kom
2. Sekretaris : Titi Sriwahyuni, S.Pd, M.Eng
3. Anggota : Muhammad Adri, S.Pd, MT
4. Anggota : Oktoria, S.Pd, MT
5. Anggota : Drs. Zulhendra, M.Kom

1. 
2. 
3. 
4. 
5. 

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa tugas akhir ini benar-benar karya saya sendiri. Sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang ditulis atau diterbitkan orang lain kecuali sebagai acuan atau kutipan dengan mengikuti tata penulisan karya ilmiah yang lazim.

Padang, September 2014

Yang menyatakan,



Marisa Rahma Danti

Kata Pengantar

Puji syukur berkat karunia Tuhan Yang Maha Esa penulis dapat menyelesaikan laporan akhir Skripsi ini yang berjudul **Penerapan Metode Analisis Asosiasi pada Data Penjualan Barang Menggunakan Algoritma Apriori (Studi Kasus pada Supermarket Robinson Plaza Andalas).**

Teriring ucapan terima kasih serta penghargaan kepada semua pihak yang telah banyak membantu dan mendukung demi rampungnya laporan akhir skripsi ini, untuk itu penulis mengucapkan terimakasih kepada :

1. Bapak Drs. Putra Jaya, MT, sebagai Ketua Jurusan Teknik Elektronika Fakultas Teknik Universitas Negeri Padang.
2. Bapak Muhammad Adri, S.Pd,MT, sebagai Pembimbing I.
3. Bapak Oktorina S.Pd,MT, sebagai Pembimbing II.
4. Bapak Drs. Zulhendra MT sebagai dosen penguji.
5. Bapak Asrul Huda, S.Kom, M.Kom sebagai dosen penguji.
6. Ibuk Titi Sriwahyuni, S.Pd, M.Eng sebagai dosen penguji.
7. Seluruh staf dosen pengajar Teknik Elektronika Fakultas Teknik Universitas Negeri Padang, yang telah mendidik dan memberikan berbagai bekal pengetahuan yang tidak ternilai harganya kepada penulis selama mengikuti perkuliahan.

8. Semua teman-teman dan pihak yang telah mendukung dalam penyusunan skripsi ini, khususnya Angkatan 2010 Teknik Elektronika Fakultas Teknik Universitas Negeri Padang.

Sebagai manusia yang tidak luput dari kesalahan, penulis mengharapkan saran dan kritik yang sifatnya membangun demi kesempurnaan laporan akhir skripsi ini. Akhir kata semoga laporan ini dapat bermanfaat. Amin.

Padang, Agustus 2014

Penulis

ABSTRAK

Marisa Rahma Danti. 2014. “ Penerapan Metode Analisis Asosiasi pada Data Penjualan Barang Menggunakan Algoritma Apriori (Studi Kasus pada Supermarket Robinson Plaza Andalas)

Penelitian ini bertujuan untuk mengolah data transaksi penjualan untuk menemukan pola tertentu yang mengasosiasikan data yang satu dengan data yang lainnya dengan menerapkan metode analisis asosiasi menggunakan algoritma apriori pada data penjualan barang. Kemudian membandingkan hasil analisis asosiasi menggunakan algoritma apriori secara manual dan dengan menggunakan *tools Data Mining* WEKA 3.7.4 dan Tanagra 1.4. Berdasarkan penelitian dengan menggunakan algoritma apriori maka diketahui bahwa algoritma apriori mengekstrak aturan berdasarkan nilai *support* dan *confidence* tinggi. Algoritma apriori mencari item frekuen pada setiap transaksi. Hasil analisis dataset penjualan barang yang terdiri dari 34 item dengan batas minimum *support* 30% menghasilkan 4 kelompok kombinasi frekuen itemset yang terdiri dari 15 itemset frekuen untuk kombinasi 1-itemset, 30 itemset frekuen untuk kombinasi 2-itemset, 19 itemset frekuen untuk kombinasi 3-itemset dan 3 itemset frekuen untuk kombinasi 4-itemset. Dari itemset yang telah ditemukan dengan algoritma apriori kemudian dibentuk aturan asosiasi dan menghitung nilai *confidence*, aturan yang tidak memenuhi nilai minimum *confidence* 70% akan dipangkas sehingga dalam penelitian ini ditemukan 130 aturan asosiasi. Dari 130 aturan asosiasi maka dicari aturan yang merupakan *Strong Rule* yaitu aturan yang memiliki nilai *support* dan *confidence* tertinggi. Hasil analisis dengan algoritma apriori menemukan 5 aturan asosiasi yang memiliki nilai *support* dan *confidence* terbesar dan dimasukkan ke dalam kelompok *Strong Rules*. Kombinasi item yang dihasilkan yang berupa aturan asosiasi dapat membantu pihak manajemen Supermarket Robinson Plaza Andalas untuk menemukan keterkaitan atau pola kemunculan barang dalam transaksi penjualan, yang pada akhirnya dapat digunakan untuk menyusun strategi penjualan.

Kata kunci : Analisis Asosiasi, Data Mining, Algoritma Apriori

DAFTAR ISI

HALAMAN JUDUL

KATA PENGANTAR..... i

ABSTRAK..... iii

DAFTAR ISI iv

DAFTAR TABEL vii

DAFTAR GAMBAR..... viii

DAFTAR LAMPIRAN ix

BAB I PENDAHULUAN

1. Latar Belakang Masalah	1
2. Identifikasi Masalah	5
3. Batasan Masalah.....	6
4. Rumusan Masalah	6
5. Tujuan Penelitian.....	7
6. Manfaat Penelitian.....	7

BAB II KAJIAN PUSTAKA

A. Data <i>Mining</i>	9
1. Tujuan data mining	13
2. Metode-metode dalam data mining	14
3. Pekerjaan dalam data mining	16
4. Data Informasi , dan Pengetahuan (Knowledge)	18
5. Set Data	19
B. Hubungan data mining dengan <i>data warehouse</i>	21
C. <i>Market Basket Analysis</i> (MBA)	23
D. Analisis asosiasi	24

1. Pengertian Asosiasi.....	24
2. Metode dasar analisis asosiasi	26
3. Defenisi-defenisi yang terdapat pada <i>association rule</i>	28
4. Algoritma-algoritma asosiasi	29
a. Algoritma apriori	29
b. Algoritma <i>Hash Based</i>	30
c. Algoritma <i>FT-Growth</i>	31
E. Algoritma Apriori.....	31
1. Konsep algoritma apriori.....	31
2. Tahapan algoritma apriori	34
F. Hubungan Market Basket Analysis dengan asosiasi pada algoritma apriori.....	36
G. Software Aplikasi Data Mining	37
1. WEKA 3.7.4 <i>Data Mining Tool</i>	37
2. Tanagra <i>Data Mining Tool</i>	38
H. Jurnal terkait.....	40
I. Gambaran Umum Manajemen Supermarket Robinson Plaza Andalas .	42

BAB III METODE PENELITIAN

A. Instrumen Penelitian.....	48
B. Desain Penelitian.....	48
C. Populasi dan Sampel Penelitian	54
D. Teknik Pengumpulan Data.....	56
E. Teknik Analisis Data	56
1. <i>Frequent Itemset Generation</i> dengan Algoritma Apriori	57
2. <i>Rule Generation</i>	62

BAB IV HASIL DAN PEMBAHASAN

A. Analisis asosiasi Data Penjualan Barang dengan Algoritma Apriori....	63
1. Praproses Data	63

a. Seleksi Data	63
b. Pembersihan Data	64
c. Transformasi Data.....	64
2. Proses Analisis Asosiasi dengan Algoritma Apriori.....	65
B. Pengujian dengan Tool Data Mining.....	71
1. WEKA 3.7.4.....	71
2. Tanagra 1.4.....	77
C. Analisis Hasil	82
D. Pembahasan.....	85

BAB V PENUTUP

A. Simpulan	87
B. Saran	88

DAFTAR PUSTAKA.....	89
----------------------------	-----------

DAFTAR TABEL

	Halaman
Tabel 1. Contoh Data Keranjang Belanja.....	21
Tabel 2. Tabel Guilford (Sugiyono 2010).....	51
Tabel 3. Perhitungan <i>support</i> masing – masing item.....	65
Tabel 4. <i>Frekuent item</i> untuk kelompok 1-itemset.....	66
Tabel 5. Perhitungan <i>support</i> untuk kandidat 4-itemset	68
Tabel 6. <i>Frekuent item</i> untuk kelompok 4-itemset.....	69
Tabel 7. Aturan asosiasi yang tergolong <i>strong rules</i>	71
Tabel 8. Jumlah frekuen itemset untuk setiap kombinasi itemset	85

DAFTAR GAMBAR

	Halaman
Gambar 1. Tahapan dalam KDD (Fajar 2013)	12
Gambar 2. Hubungan Data <i>Warehouse</i> dan Data <i>Mining</i>	23
Gambar 3. Hubungan MBA dengan Asosiasi dan Algoritma Apriori.....	37
Gambar 4. Desain Penelitian	49
Gambar 5. Flowchart algoritma apriori.....	58
Gambar 6. Ilustrasi algoritma Apriori.....	60
Gambar 7. Data dalam format *.xlsx.....	71
Gambar 8. Data dalam format *.arff.....	72
Gambar 9. Tampilan WEKA GUI Chooser	72
Gambar 10. Tampilan WEKA explorer setelah data dimasukkan.....	73
Gambar 11. Tampilan weka explorer untuk memilih algoritma yang digunakan dan mengubah parameter	74
Gambar 12. Output WEKA (<i>Run Information</i>)	75
Gambar 13. Output WEKA (parameter dan jumlah itemset)	75
Gambar 14. Hasil output WEKA (Aturan Asosiasi).....	76
Gambar 15. Dataset dengan format *.arff.....	77
Gambar 16. Input dataset ke Tanagra	78
Gambar 17. Define status untuk input variable	79
Gambar 18. Pengaturan parameter	80
Gambar 19. Hasil output Tanagra menampilkan parameter	80
Gambar 20. Hasil output Tanagra (<i>item section</i>).....	81
Gambar 21. Hasil output Tanagra berupa aturan asosiasi (<i>Rules</i>).....	82

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Dataset penjualan	91
Lampiran 2. Perhitungan <i>support</i> untuk 1-itemset	95
Lampiran 3. Pembangkitan kandidat 2-itemset	98
Lampiran 4. Perhitungan nilai <i>support</i> kandidat 2-itemset	99
Lampiran 5. Frekuen item untuk kandidat 2-itemset	104
Lampiran 6. Pembangkitan kandidat 3-itemset	106
Lampiran 7. Perhitungan nilai <i>support</i> kandidat 3-itemset	109
Lampiran 8. <i>Frequent itemset</i> untuk kelompok 3-itemset	111
Lampiran 9. Pembangkitan kandidat 4-itemset	113
Lampiran 10. Pembangkitan kandidat 5-itemset	118
Lampiran 11. Kandidat aturan asosiasi	119
Lampiran 12. Aturan asosiasi yang dihasilkan	135
Lampiran 13. Data penelitian dari Robinson Plaza Andalas	144
Lampiran 14. Surat izin penelitian dari Jurusan Teknik Elektronika	156
Lampiran 15. Surat izin penelitian dari Fakultas Teknik	157
Lampiran 16. Surat pelaksanaan penelitian dari Robinson Plaza Andalas	158

BAB I

PENDAHULUAN

A. Latar belakang

Setiap harinya sebuah supermarket mengakumulasi data transaksi dalam jumlah besar. Jika dalam satu hari ada 100 transaksi, dalam setahun setidaknya ada 36.000 transaksi, bisa dibayangkan jumlahnya jika sudah bertahun-tahun. Kemudian setelah data ini selesai digunakan setiap tahunnya, untuk apa data tersebut? Apakah dibuang? Apakah disimpan begitu saja hingga menggunung jumlahnya? Tentu saja, meskipun hanya disimpan sebagai arsip, ada biaya yang harus dikeluarkan perusahaan untuk pemeliharaannya. Jika data-data yang sudah terakumulasi itu tidak dimanfaatkan, padahal ada biaya perawatan yang harus dibayar, perusahaan akan mengalami kerugian.

Robinson Plaza Andalas merupakan salah satu contoh perusahaan ritel modern yang cukup berkembang di Kota Padang saat ini. Berbagai macam kebutuhan konsumen disediakan pada ritel ini. Untuk menarik konsumennya, Robinson harus menerapkan strategi pemasaran dengan baik. Selain itu perusahaan ini harus bisa memperhatikan kepuasan pelanggan akan faktor-faktor pemasaran yang diterapkan, karena kepuasan konsumen merupakan tujuan utama perusahaan.

Sejak berdiri sampai sekarang, Robinson telah menjadi salah satu pusat perbelanjaan ternama di kota Padang. Ini dibuktikan dengan besarnya jumlah pelanggan yang terlihat dari besarnya jumlah transaksi tiap bulan.

Supermarket Robinson Plaza Andalas setiap harinya mengakumulasi data penjualan yang begitu banyak. Terdapat sekitar 20.000 item yang dijual di Supermarket Robinson Plaza Andalas dan untuk item makanan dan kebutuhan sehari-hari ada 8672 item. Banyaknya item yang dijual tentunya akan sulit untuk mengatur tata letak barang dan mengetahui pola pembelian konsumen.

inventori (stok barang) juga merupakan permasalahan operasional yang sering dihadapi oleh supermarket. Jika jumlah inventori terlalu sedikit dan permintaan tidak dapat dipenuhi karena kekurangan persediaan, maka akan mengakibatkan konsumen merasa kecewa dan ada kemungkinan konsumen tidak akan kembali lagi. Begitu juga jika inventori terlalu besar, maka akan mengakibatkan kerugian bagi supermarket karena harus menyediakan tempat yang lebih besar, terjadinya penyusutan nilai guna barang, serta harus menyediakan biaya tambahan yang terkait dengan biaya inventori seperti biaya pemeliharaan dan biaya akuntansi.

Oleh sebab itu pihak manajemen supermarket perlu mengetahui pola pembelian pelanggan. Salah satu cara yang bisa dilakukan untuk mengetahui kondisi pasar dan pola pembelian pelanggan adalah dengan mengamati data keranjang belanja pembeli di sebuah supermarket.

Data transaksi penjualan yang terkumpul dan tersimpan di *database* dapat memberikan pengetahuan yang bermanfaat bagi pengambil keputusan dalam melakukan usaha-usaha yang terkait dengan peningkatan penjualan

dengan melakukan promosi yang tepat dan membekali tenaga penjual dengan pengetahuan mengenai kebiasaan customer.

Pengetahuan mengenai keterkaitan antar jenis item, jenis-jenis item yang sering muncul pada tiap transaksi, dapat menjadi masukan penting dalam melakukan usaha peningkatan penjualan. Dengan mengetahui pola pembelian pelanggan, manajemen supermarket dapat membuat keputusan, misalnya, bagaimana strategi untuk menghabiskan barang yang kurang laku, dan kapan waktu yang tepat untuk promosi diskon barang, pengetahuan asosiasi antar jenis item ini dapat memberikan pengetahuan mengenai kombinasi item yang harus tercakup dalam promosi tersebut.

Permasalahan yang muncul adalah bagaimana memproses data-data transaksi penjualan tersebut sehingga dapat menampilkan suatu pengetahuan yang berguna bagi manajemen dalam mengambil keputusan. Namun seiring dengan perkembangan Teknologi Informasi (TI) muncul berbagai cara dan solusi untuk mengatasi dan meminimalisir kesulitan dalam mengidentifikasi data penjualan tersebut, diantaranya dengan menggunakan metode pada teknik *Data Mining* (DM). Metode yang digunakan yaitu metode asosiasi, dengan metode ini dimaksudkan dapat mengetahui keterkaitan antar item pada data transaksi penjualan dalam wujud aturan asosiasi.

Eko (2012: 1) mengemukakan bahwa "*Data mining* adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan.

Selain itu menurut Santoso (2007) dalam Ong (2013:13) “*Data mining* merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, *clustering*, klasifikasi, regresi, seleksi variabel, dan asosiasi”.

Berdasarkan pendapat ahli di atas dapat disimpulkan data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Salah satu metode yang seringkali digunakan dalam teknologi data mining adalah metode asosiasi atau *association rule mining*. Di dalam bidang usaha retail metode *association rule mining* ini lebih dikenal dengan istilah analisa keranjang belanja (*market basket analysis*). *Market basket analysis* adalah suatu metode analisa atas perilaku konsumen secara spesifik dari suatu golongan / kelompok tertentu. Metode asosiasi 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*). Yang dimaksud dengan *association rules* adalah pola-pola keterkaitan data dalam basis data.

Dalam penelitian ini penulis akan mencoba untuk melakukan suatu eksperimen terhadap data transaksi penjualan barang di Supermarket Robinson Plaza Andalas dengan menggunakan sebuah metode data mining

analisis asosiasi untuk melihat hubungan asosiasi (korelasi) antara sejumlah atribut penjualan.

Terdapat banyak algoritma yang dapat digunakan pada teknik *data mining* dengan metode asosiasi, salah satunya adalah algoritma apriori. Algoritma Apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule* (Erwin, 2009).

Diterapkannya algoritma apriori pada penelitian ini diharapkan dapat memberikan suatu pengetahuan baru tentang hubungan asosiasi pada sejumlah atribut penjualan sehingga dapat meningkatkan hasil penjualan dan strategi pemasaran serta promosi pada Supermarket Robinson Plaza Andalas.

Berdasarkan permasalahan tersebut diatas maka penulis memberi judul skripsi **“Penerapan Metode Analisis Asosiasi pada Data Penjualan Barang Menggunakan Algoritma Apriori (studi kasus pada Supermarket Robinson Plaza Andalas)”**.

B. Identifikasi Masalah

Berdasarkan latar belakang yang telah dipaparkan diidentifikasi masalahnya adalah sebagai berikut :

1. Menerapkan metode analisis asosiasi dengan algoritma apriori pada data penjualan barang di Supermarket Robinson Plaza Andalas.
2. Cara pemanfaatan dan pengelolaan data transaksi penjualan dengan metode asosiasi menggunakan algoritma apriori sehingga dapat diketahui pola pembelian pelanggan.

3. Analisis asosiasi dengan algoritma apriori untuk mengolah data penjualan sehingga bisa menemukan pola tertentu yang mengasosiasikan data yang satu dengan data yang lain.
4. Bagaimana mencari aturan asosiatif atau aturan keterkaitan antar setiap atribut penjualan tersebut.

C. Batasan Masalah

Berdasarkan identifikasi masalah, maka batasan masalah dalam menganalisis data penjualan barang menggunakan algoritma apriori adalah :

1. Analisis data penjualan barang dengan metode asosiasi menggunakan algoritma apriori.
2. Data yang digunakan adalah data transaksi penjualan 1 tahun terakhir periode Mei 2013 – April 2014 yang diperoleh dari Supermarket Robinson Plaza Andalas. Data penjualan yang digunakan adalah data penjualan perbulan.
3. Atribut yang digunakan dalam dataset ini adalah bulan dan nama barang/item.
4. Tool Data mining yang digunakan adalah WEKA 3.7.4 dan Tanagra 1.4.

D. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dipaparkan diatas, dapat dirumuskan masalah yang diteliti adalah Bagaimana menerapkan metode analisis asosiasi dengan algoritma apriori pada data penjualan barang di

Supermarket Robinson Plaza Andalas untuk mengetahui pola pembelian pelanggan dalam bentuk aturan asosiasi yang dihasilkan.

E. Tujuan Penelitian

Sesuai dengan latar belakang, batasan masalah serta rumusan masalah diatas, maka penelitian ini bertujuan untuk :

1. Menerapkan metode analisis asosiasi menggunakan algoritma apriori pada data penjualan barang.
2. Mengolah data transaksi penjualan untuk menemukan pola tertentu yang mengasosiasikan data yang satu dengan data yang lain.
3. Mengetahui pola transaksi penjualan selama satu tahun terakhir di Supermarket Robinson Plaza Andalas.
4. Membandingkan hasil analisis asosiasi menggunakan algoritma apriori secara manual dan dengan menggunakan Tool data mining WEKA 3.7.4 dan Tanagra 1.4.

F. Manfaat Penelitian

Manfaat yang diperoleh dalam penelitian ini yaitu :

1. Pihak pengelola supermarket dapat mengetahui pola transaksi penjualan selama satu tahun terakhir di Supermarket Robinson Plaza Andalas.
2. Memberikan pengetahuan tentang langkah – langkah analisis asosiasi menggunakan algoritma Apriori.
3. Memperoleh pengetahuan mengenai asosiasi antar item sehingga dapat bermanfaat bagi pengambil keputusan dalam melakukan usaha-usaha

yang terkait dengan peningkatan penjualan dengan melakukan promosi yang tepat, menentukan jumlah stok barang yang harus disediakan dan membekali tenaga penjual dengan pengetahuan mengenai kebiasaan *customer*.

BAB II

KAJIAN TEORI

A. Data Mining

Menurut Eko (2012: 3) “*Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial yang bermanfaat dan tersimpan didalam database besar”.

Eko (2012: 1) mengemukakan bahwa “*Data mining* adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan.

Larose (2005) dalam Prabowo, Rahmadya, dan Herlawati (2013:3) menyatakan bahwa “*data mining* adalah suatu proses pencarian korelasi, pola dan tren baru yang berguna dalam media penyimpanan data berukuran besar menggunakan teknologi pengenalan pola seperti teknik-teknik statistik dan matematis. Istilah lain yang sering digunakan antara lain *knowledge mining from data*, *knowledge extraction*, *data/pattern analysis*, *data archeology*, dan *data dredging*”.

Selain itu menurut Santoso (2007) dalam Johan (2013:13) “*Data mining* merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan

teknologi. Secara umum kajian data mining membahas metode-metode seperti, *clustering*, klasifikasi, regresi, seleksi *variabel*, dan market basket analisis”.

Menurut Sani dan Dedy (2010:2) “*Data mining* merupakan disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki”. Hal ini sejalan dengan pendapat Larose (2006) bahwa “*Data mining* juga merupakan bidang dari beberapa keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar”.

Fajar (2013:3) “*Data mining* merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang sah (sempurna), bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar (*massive database*)”.

Data Mining adalah sebuah teknologi baru yang memiliki potensi sangat besar dalam penggalian informasi yang tersembunyi. *Data mining* juga sering didefinisikan sebagai proses ekstraksi informasi prediktif tersembunyi dari database yang sangat besar. Teknologi ini memungkinkan sebuah perusahaan untuk lebih proaktif dalam penyusunan strategi yang efektif dan juga dalam pengambilan keputusan berdasarkan perilaku dan tren masa depan yang dapat diprediksi oleh proses *data mining*.

Data mining membantu perusahaan untuk mendapatkan pola dari data-data yang tersimpan di dalam basis data perusahaan. Pengetahuan yang diperoleh tersebut akan menjadi pedoman dalam mengambil tindakan-tindakan

bisnis sebagai upaya pemeliharaan dan peningkatan tingkat kompetitif bisnis perusahaan. Walaupun sudah banyak perangkat lunak yang menawarkan kemampuan dalam proses *data mining*, keterlibatan manusia sangat dibutuhkan dalam setiap fase proses *data mining* itu sendiri. Pemahaman terhadap model statistik dan matematik yang digunakan dalam perangkat lunak sangat dituntut.

Berdasarkan pendapat ahli di atas dapat disimpulkan data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.

Berdasarkan definisi-definisi yang telah disampaikan, hal penting yang terkait dengan *data mining* adalah :

1. *Data mining* merupakan suatu proses penambangan, menggali pengetahuan terhadap data yang sudah ada.
2. Data yang akan diproses berupa data berskala besar.
3. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Istilah *data mining* dan *Knowledge Discovery in Database* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu kumpulan data yang besar.

Fajar (2013:3) menyatakan *Knowledge Discovery in Databases* (KDD) adalah penerapan metode saintifik pada *data mining*. Dalam konteks ini *data mining* merupakan satu langkah dari proses KDD.

KDD terdiri dari tiga proses utama yaitu:

1. *Preprocessing*

Preprocessing dilakukan terhadap data sebelum algoritma *data mining* diaplikasikan. Proses ini meliputi *data cleaning*, integrasi, seleksi dan transformasi.

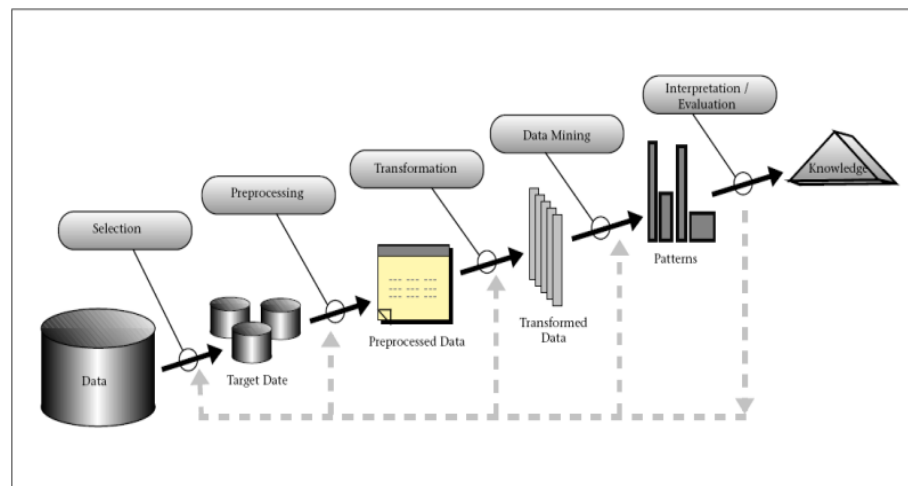
2. *Data mining*

Proses utama dalam KDD adalah proses *data mining*, dalam proses ini algoritma-algoritma *data mining* diaplikasikan untuk mendapatkan pengetahuan dari sumber data.

3. *Post processing*

Hasil yang diperoleh dari proses *data mining* selanjutnya akan dievaluasi pada *post processing*.

Tahapan proses dalam menggunakan data mining yang merupakan proses *Knowledge Discovery in Database (KDD)* seperti terlihat pada Gambar 1.



Gambar 1 Tahapan dalam KDD (Fajar 2013)

Mengacu pada Fajar (2013:5) tahapan proses *Knowledge Discovery in Database* (KDD) dapat diuraikan sebagai berikut :

1. Memahami domain aplikasi untuk mengetahui dan menggali pengetahuan awal serta apa sasaran pengguna.
2. Membuat target data-set yang meliputi pemilihan data dan focus pada subset data.
3. Pembersihan dan transformasi data meliputi eliminasi derau, *outliers*, *missing value* serta pemilihan fitur dan reduksi dimensi.
4. Penggunaan algoritma data mining
5. Interpretasi, evaluasi dan visualisasi pola untuk melihat apakah ada sesuatu yang baru dan menarik dan dilakukan iterasi jika diperlukan.

1. Tujuan Data Mining

Baskoro (2010) dalam Eko (2012:55) menyatakan bahwa tujuan dari data *mining* adalah :

- a. *Explanatory*, yaitu untuk menjelaskan beberapa kegiatan observasi atau suatu kondisi.
- b. *Confirmatory*, yaitu untuk mengkonfirmasi suatu hipotesis yang telah ada.
- c. *Exploratory*, yaitu untuk menganalisis data baru suatu relasi yang janggal.

2. Metode - Metode Dalam Data Mining

Mengacu kepada Larose (2005), terdapat enam metode data mining, yaitu

a. Metode Deskripsi (*deskription*)

Pola dan trend data sering dideskripsikan. Ketika diberi sekumpulan data, terkadang agak sukar bagi kita untuk menangkap arti kumpulan data tersebut. Sekumpulan data tersebut perlu dirangkum sedemikian rupa agar dapat menghasilkan pengetahuan sehingga kita memiliki gambaran mengenai kumpulan data tersebut. Deskripsi tersebut sangat membantu dalam menjelaskan pola dan trend yang terjadi. Model *data mining* harus setransparan mungkin, dimana hasilnya dapat mendeskripsikan pola dengan jelas.

b. Metode Estimasi (*estimation*)

Estimasi adalah fungsi untuk memperkirakan suatu kumpulan data sehingga menghasilkan suatu pengetahuan. Estimasi mirip dengan klasifikasi kecuali variabel target-nya numerik ketimbang kategorikal. Model yang dibangun menggunakan *record* yang lengkap, yang menyediakan nilai variabel target dan *predictor*. Untuk observasi yang baru, estimasi nilai variabel target ditentukan, berdasarkan nilai-nilai *predictor*.

c. Metode Prediksi (*prediction*)

Prediksi mirip dengan klasifikasi dan estimasi, dalam prediksi kita menggunakan data yang ada untuk memprediksi hasil dari suatu hal baru yang akan muncul selanjutnya.

d. Metode Klasifikasi (*classification*)

Dalam klasifikasi, variabel target-nya merupakan kategorikal. Model *data mining* memeriksa set *record* yang besar, tiap *record* mempunyai informasi variabel target dan set input atau variabel *predictor*. Pada klasifikasi, sebuah catatan (*record*) akan diklasifikasikan kedalam salah satu dari sekian klasifikasi yang tersedia pada variabel tujuan berdasarkan nilai-nilai variabel *predictor*nya.

e. Metode Pengelompokan (*clustering*)

Clustering merupakan pengelompokan *record*, observasi, atau kasus ke dalam kelas-kelas objek yang mirip. *Clustering* berbeda dengan klasifikasi dimana dalam *clustering* tidak terdapat variabel target. *Clustering* mencoba menyegmentasi seluruh set data ke dalam *subgroup* atau *cluster* yang relatif homogen, dimana kemiripan antar *record* dalam *cluster* dimaksimalkan dan kemiripan *record* di luar *cluster* diminimasi.

f. Metode Asosiasi (*association*)

Asosiasi merupakan suatu tugas untuk menemukan atribut-atribut yang “terjadi” bersamaan. Tugas asosiasi mencoba untuk menemukan aturan untuk mengkuantifikasi hubungan antara dua atau lebih atribut. Aturan asosiasi berbentuk “If *antecedent*, then *consequent*”, bersama-sama dengan ukuran *support* dan *confidence* yang berhubungan dengan aturan. Tugas asosiasi dalam *data mining* adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari

asosiasi adalah *market basket analysis* atau analisis keranjang belanja, sebagaimana yang akan dibahas dalam penelitian ini.

Mengacu pada Berry dan Browne (2006) dalam Dedy dan Sani (2010:5), keenam fungsi data mining tersebut dapat dipilah menjadi:

- a. Fungsi minor atau fungsi tambahan, yaitu deskripsi, estimasi, dan prediksi.
- b. Fungsi mayor atau fungsi utama, yaitu klasifikasi, pengelompokan, dan asosiasi.

3. Pekerjaan Dalam *Data Mining*

Eko (2012:5) menyatakan pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi 4 kelompok, yaitu model prediksi, analisis kelompok, analisis asosiasi, dan deteksi anomaly.

1. Model prediksi

Model prediksi berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada dua jenis model prediksi yaitu klasifikasi dan regresi.

2. Analisis kelompok

Analisis kelompok melakukan pengelompokan data-data kedalam sejumlah kelompok (cluster) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada.

Data-data yang masuk dalam batas kesamaan dalam kelompoknya akan bergabung dalam kelompok tersebut.

3. Analisis asosiasi

Analisis asosiasi digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya merepresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien.

Penerapan yang paling dekat dengan kehidupan sehari-hari adalah analisa keranjang pasar. Sebagai contoh, pembeli adalah ibu rumah tangga yang akan membeli barang kebutuhan rumah tangga di supermarket. Jika ibu membeli beras, sangat besar kemungkinannya bahwa ibu itu juga akan membeli barang lain, misalnya minyak, telur dan tidak mungkin atau jarang membeli barang lain seperti topi atau buku. Dengan mengetahui hubungan yang lebih kuat antara beras dengan telur daripada beras dengan topi, pengencer dapat menentukan barang-barang yang sebaiknya disediakan dalam jumlah yang cukup banyak.

4. Deteksi anomaly

Pekerjaan deteksi anomaly berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain. Deteksi anomaly dapat diterapkan pada system jaringan untuk mengetahui pola data

yang memasuki jaringan sehingga penyusupan bisa ditemukan jika pola kerja data yang dating berbeda.

4. Data, Informasi, dan Pengetahuan (*Knowledge*)

Menurut Eko (2012:4) Data adalah segala fakta, angka, atau teks yang dapat diproses oleh komputer. Data-data tersebut dibagi menjadi 3 jenis :

1. Data operasional atau data transaksional, seperti penjualan, inventaris, penggajian, akuntansi dan sebagainya.
2. Data non operasional, seperti industry penjualan (supermarket), peramalan, dan data ekonomi makro.
3. Metadata, adalah data mengenai data itu sendiri, seperti desain logika basis data atau defenisi kamus data.

Sementara, informasi adalah pola, asosiasi, atau hubungan antara semua data yang dapat memberikan informasi. Sebagai contoh, analisis titik eceran data transaksi penjualan dapat menghasilkan informasi mengenai produk apa yang sebaiknya dijual dan kapan menjualnya.

Informasi dapat dikonversi menjadi pengetahuan mengenai pola-pola historis dan tren masa depan. Misalnya, ringkasan informasi tentang penjualan eceran supermarket dapat dianalisis sehubungan dengan upaya promosi untuk memberikan pengetahuan mengenai perilaku konsumen daalam membeli. Dengan demikian, produsen atau pengecer dapat menentukan item yang paling rentan terhadap upaya promosi.

5. Set Data

Eko (2012;9) mengemukakan bahwa set data (data set) dapat dipandang sebagai kumpulan objek data. Nama lainnya adalah *record*, *point*, *vector*, *pattern*, *event*, *observation*, *case*, atau data. Objek data digambarkan dengan sejumlah atribut yang menangkap (*capture*) karakter dasar objek data. Atribut juga disebut variabel, karakteristik, medan (*field*), fitur, atau dimensi“.

Atribut adalah sifat atau properti atau karakteristik objek data yang nilainya bisa bermacam-macam dari satu objek ke objek yang lain, dari satu waktu ke waktu yang lain. Atribut yang menjadi elemen setiap data mempunyai jenis yang beragam.

Fajar (2013:24-26)

Atribut dapat dibedakan dalam tipe-tipe yang berbeda bergantung pada tipe domainnya, yaitu bergantung pada tipe nilai yang diterima. Atribut kategorikal (*categorical attribute*) adalah salah satu tipe yang domainnya merupakan sebuah himpunan simbol berhingga. Atribut kategorikal dibedakan menjadi dua tipe, yaitu :

1. Nominal: sebuah atribut dikatakan nominal jika nilai-nilainya tidak dapat diurutkan.
 2. Ordinal: jika nilai-nilainya dapat diurutkan dalam beberapa cara.
- Tipe atribut kedua adalah atribut numerik (*numeric attribute*) yang domainnya berupa bilangan riil atau *integer*. Atribut numerik juga dibedakan menjadi dua, yaitu :
1. Interval: untuk jenis atribut ini mempunyai sifat bahwa perbedaan antara nilai-nilainya sangat berarti.
 2. Rasio: dalam atribut jenis ini, baik beda maupun rasio sangat berarti.

Atribut berdasarkan jumlah nilainya dibedakan menjadi dua yaitu :

1. Atribut Diskrit (*Discrete Attribute*) yaitu atribut yang hanya menggunakan sebuah himpunan nilai berhingga atau himpunan nilai tak berhingga yang dapat dihitung.
2. Atribut Kontinyu (*Continuous Attribute*) yaitu atribut yang menggunakan bilangan riil sebagai nilai atribut.

Himpunan data (data-set) mempunyai beberapa karakteristik umum yaitu :

1. Dimensionality
 - a. Dimensionalitas dari sebuah data-set adalah jumlah atribut yang dimiliki oleh objek-objek dalam data-set.
 - b. Data dengan jumlah dimensi kecil punya kecenderungan berbeda secara kualitatif dibandingkan dengan data dimensi tinggi.
 - c. Kesulitan yang berhubungan dengan data dimensi tinggi sering disebut sebagai *curse of dimensionality*.
 - d. Untuk itu pada tahap preprocessing perlu dilakukan pengurangan dimensi (*dimensionality reduction*).
2. Sparsity
 - a. Untuk beberapa *data-set*, misalkan data dengan fitur asimetris, kebanyakan atribut dari suatu objek mempunyai nilai 0; dan biasanya kurang dari 1% mempunyai nilai tidak nol.
 - b. *Sparsity* mempunyai keuntungan dalam waktu komputasi dan penyimpanan data.
3. Resolution
 - a. Sifat dari data berbeda pada resolusi yang berbeda.
 - b. Pola dalam data bergantung pada level resolusi.
 - c. Jika resolusi terlalu baik (tidak ada perbedaan/ halus), pola mungkin tidak akan kelihatan, jika resolusi terlalu kasar, pola juga akan hilang.

Kebanyakan metode data mining mengasumsikan bahwa set data yang diproses adalah kumpulan basis data (*records/ entries/ objects*), dimana setiap barisnya terdiri atas sejumlah fitur (atribut) yang tetap. Dalam set data berbentuk data record, tidak ada hubungan antara baris data yang satu dengan baris data yang lain, dan juga dengan set data yang lain. Set data yang diolah dalam data mining adalah keluaran dari system *data warehouse* yang menggunakan query untuk melakukan pengambilan data dari sejumlah table dalam system basis data.

Data keranjang belanja adalah jenis khusus dari data record, dimana setiap record (transaksi) berisi sejumlah item, dan jumlah item untuk sebuah transaksi bisa berbeda dengan transaksi yang lain.

Contohnya bias dilihat pada kasus keranjang belanja di supermarket, setiap pembeli melakukan pembelian barang yang jumlah dan jenisnya bisa berbeda dengan pembeli yang lain. Data transaksinya berisi kumpulan item, tetapi dapat dilihat bahwa field atributnya asimetrik. Biasanya atribut dengan nilai 1 untuk adanya barang (atribut) yang dibeli dilambangkan dengan biner, dan 0 untuk barang yang tidak dibeli. Contoh data keranjang belanja ditunjukkan pada Tabel 1.

Table 1 Contoh Data Keranjang Belanja

IDT	Item
1	Susu, Bedak, Sabun
2	Roti, Susu, Mentega
3	Gula, Roti, Terigu
4	Bedak, Gula, Sabun, Roti
5	Roti, Terigu

B. Hubungan Data Mining dengan *Data Warehouse*

Data warehouse merupakan penyimpanan data yang berorientasi objek, terintegrasi, mempunyai variant waktu, dan penyimpanan data berbentuk *nonvolatile* sebagai pendukung manajemen dalam proses pengambilan keputusan (Han & Kember, 2006:136).

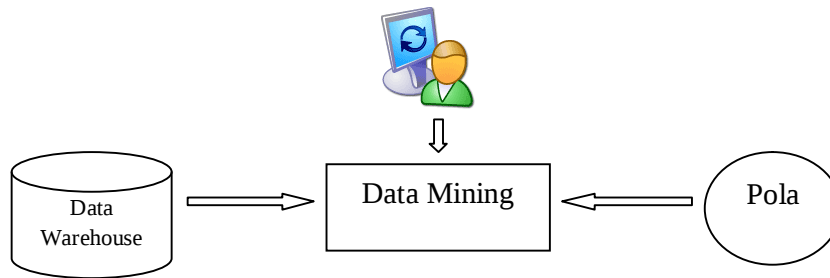
Menurut Vincent Rainardi yang menulis buku “*Building a Data Warehouse With Examples in SQL Server*” dalam Feri & Dominikus (2010:33), menyampaikan bahwa *data warehouse* merupakan suatu sistem yang

mengkonsolidasikan data secara periodik dari sistem-sistem yang ada (OLTP) ke dalam suatu penyimpanan dimensional. Pada umumnya, data *warehouse* menyimpan data histori beberapa tahun dan akan dilakukan *query* untuk keperluan organisasi atau aktifitas analisis lainnya.

Data mining berpotensi tinggi jika data yang tepat dikumpulkan dan disimpan dalam sebuah gudang data (data warehouse). Fajar (2013:6) menyatakan bahwa sebuah gudang data merupakan suatu *system managemen basisdata relasional* (RDMS) yang didesain khusus untuk memenuhi kebutuhan akan system pengolahan transaksi. Data warehouse baik untuk mengintegrasikan keseluruhan data sebuah perusahaan, tanpa memperhatikan lokasi, format, atau kebutuhan komunikasi yang memungkinkan untuk memasukkan informasi tambahan.

Gudang data (data warehouse) biasanya berisi data sejarah, terkumpul dari sumber yang berbeda-beda seperti system proses transsaksi online (OLTP), system warisan, file-file teks dan *spreadsheets*. Pada data tersebut kemudian dilakukan proses pembersihan untuk akurasi dan konsistensi dan mengelolanya untuk memudahkan dan efisiensi pada *query*.

Data mining merupakan suatu proses yang interaktif atau terotomatisasi untuk menemukan pola (*pattern*) data tersebut dan memprediksi kelakuan (*trend*) dimasa mendatang berdasarkan pola data tersebut. Bisa disimpulkan bahwa *data mining* mengambil data yang berasal dari *warehouse*, untuk kemudian nanti dilakukan pengelolaan untuk menghasilkan sebuah pola yang baik dan menguntungkan.



Gambar 2 Hubungan Data Warehouse dan Data Mining

C. Market Basket Analysis (MBA)

Market Basket Analysis adalah salah satu teknik pemodelan dalam *data mining* berdasarkan teori yang mana jika anda membeli suatu grup *item*, anda akan memiliki kemungkinan membeli *itemset* yang lain (*data mining concept and technique*, Jiawei Han).

Fungsi *association rule* seringkali disebut juga dengan analisis keranjang belanja (*Market Basket Analysis*) yang digunakan untuk menemukan relasi atau korelasi diantara himpunan item-item. Analisis keranjang belanja merupakan analisis dari kebiasaan membeli customer dengan mencari asosiasi dan korelasi antara item-item berbeda yang diletakan customer dalam keranjang belanjanya.

Menurut Olson dan Shi (2008) dalam Dedy dan Sani (2010:35), menyatakan bahwa Analisis Keranjang Belanja mengacu pada berbagai metodologi yang mempelajari komposisi keranjang belanja yang terdiri atas produk-produk yang dibeli pada kejadian belanja.

Proses ini menganalisis *buying habits* dari para konsumen dengan menemukan hubungan asosiasi antar *item-item* yang berbeda yang seringkali dibeli oleh konsumen. Hasil dari proses analisis ini nantinya akan sangat

berguna bagi perusahaan retail khususnya seperti toko swalayan dan supermarket untuk mengembangkan strategi pemasaran dan proses pengambilan keputusan dengan melihat *item-item* berbeda yang sering dibeli oleh konsumen dalam satu waktu (Han & Kember, 2006).

Masalah-masalah seperti kehabisan stok akan diminimalisir dengan diketahuinya pola pembelian konsumen melalui *Market Basket Analysis* sehingga dapat meningkatkan penjualan perusahaan tersebut.

D. Analisis Asosiasi

1. Pengertian asosiasi

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. Di dalam bidang usaha retail metode *association rule mining* ini lebih dikenal dengan istilah analisa keranjang belanja (*market basket analysis*).

Menurut Eko (2012:312) analisis asosiasi berguna untuk menemukan hubungan penting yang tersembunyi di antara set data yang sangat besar. Hubungan yang sudah terbuka direpresentasikan dalam bentuk aturan asosiasi atau set aturan item yang sering muncul.

Tujuan analisis asosiasi adalah menemukan pola tertentu dari data yang berjumlah sangat besar.

Association rule adalah salah satu teknik utama atau prosedur dalam *Market Basket Analysis* untuk mencari hubungan antar item dalam suatu *data set* dan menampilkan dalam bentuk *association rule* (Budhi dkk, 2007). *Association rule* (aturan asosiatif) akan menemukan pola

tertentu yang mengasosiasikan data yang satu dengan data yang lain. Untuk mencari *association rule* dari suatu kumpulan data, tahap pertama yang harus dilakukan adalah mencari *frequent itemset* terlebih dahulu. *Frequent itemset* adalah sekumpulan item yang sering muncul pada transaksi penjualan. Setelah semua pola *frequent itemset* ditemukan, barulah mencari aturan asosiatif atau aturan keterkaitan yang memenuhi syarat yang telah ditentukan.

Jika diasumsikan bahwa barang yang dijual di swalayan adalah semesta, maka setiap barang akan memiliki *boolean* variabel yang akan menunjukkan keberadaannya atau tidak barang tersebut dalam satu transaksi atau satu keranjang belanja. Pola *boolean* yang didapat digunakan untuk menganalisa barang yang sering dibeli. Pola tersebut dapat dirumuskan dalam sebuah *association rule*. Sebagai contoh konsumen biasanya akan membeli jeruk dan tomat yang ditunjukkan sebagai berikut :

Jeruk $\rightarrow$ Tomat [*support* =2%, *confidence*=60%]

Association rule diperlukan suatu variabel ukuran yang ditentukan sendiri oleh *user* untuk menentukan batasan sejauh mana atau sebanyak apa *output* yang diinginkan *user*.

Support dan *confidence* adalah *sebuah* ukuran kepercayaan dan kegunaan suatu pola yang telah ditemukan. Nilai *support* 2% menunjukkan bahwa keseluruhan dari total transaksi konsumen membeli jeruk dan tomat yaitu sebanyak 2%. Sedangkan *confidence* 60% yaitu menunjukkan bila konsumen membeli jeruk dan pasti membeli tomat sebesar 60%.

2. Metode dasar analisis asosiasi

Interestingness measure yang dapat digunakan dalam *data mining* adalah :

1. *Support*, adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *itemset* dari keseluruhan transaksi.

Nilai *support* sebuah item diperoleh dengan rumus berikut:

$$\text{support}(A) = \frac{\text{jumlah transaksi mengandung } A}{\text{total transaksi}} \times 100\%$$

sedangkan nilai *support* dari 2 item diperoleh dari rumus berikut:

$$\text{support}(A \cap B) = \frac{\text{jumlah transaksi mengandung } A \text{ dan } B}{\text{total transaksi}} \times 100\%$$

2. *Confidence*, adalah suatu ukuran yang menunjukkan hubungan antar dua *item* secara *conditional* (berdasarkan suatu kondisi tertentu).

Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut:

$$\text{Confidence} = P(B | A) = \frac{\text{jumlah transaksi mengandung } A \text{ dan } B}{\text{jumlah transaksi mengandung } A} \times 100\%$$

Metode dasar analisis asosiasi dibagi dua tahap :

1. Analisis frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam database.

2. Pembentukan aturan asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$.

Menurut Fajar (2013:101) dalam menentukan *Association Rule Mining* dapat digunakan pendekatan dua langkah, yaitu :

1. *FrequentItemset Generation*

Membangkitkan semua itemset yang support ³ minsup

2. *Rule Generation*

Membangkitkan *rule* dengan *confidence* tinggi dari setiap *FrequentItemset*, dimana setiap *rule* merupakan partisi biner dari semua *FrequentItemset*

Sejalan dengan pendapat Eko (2012:315) strategi umum yang diadopsi oleh banyak algoritma penggalian aturan asosiasi adalah memecahkan masalah kedalam dua pekerjaan utama, yaitu:

1. *Frequent Itemset Generation*

Tujuannya adalah untuk mencari semua itemset yang memenuhi ambang batas minsup. Itemset itu disebut item frequent (itemset yang sering muncul).

2. *Rule Generation*

Tujuannya adalah mengekstrak aturan dengan *confidence* tinggi dari itemset frequen yang ditemukan dalam langkah sebelumnya. Aturan ini kemudian disebut aturan yang kuat (*strong rule*)

Dari beberapa pendapat diatas dapat disimpulkan bahwa dalam analisis asosiasi dengan algoritma apriori memiliki 2 tahapan. Tahap pertama algoritma apriori adalah menentukan itemset yang memenuhi batas *minimum support*, pada tahap inikita membangkitkan seluruh itemset yang

frekuen sehingga tidak ada lagi itemset yang bisa dibangkitkan. Pembangkitan itemset berakhir ketika tidak ada lagi kombinasi itemset yang memenuhi batas *minimum support*. Langkah kedua adalah membentuk aturan asosiasi dari itemset frekuen yang ditemukan dengan menghitung nilai *confidence* dari setiap kombinasi itemset. Itemset yang memenuhi nilai batas minimum *confidence* dan nilai *confidence* tinggi akan menjadi *strong rules*.

3. Definisi-definisi yang terdapat pada *Association Rule*

1. I adalah himpunan yang tengah dibicarakan.

Contoh: {A,B,C,....F}

2. D adalah Himpunan seluruh transaksi yang tengah dibicarakan

Contoh: {Transaksi 1, transaksi 2, ..., transaksi 14}

3. Proper Subset adalah Himpunan Bagian murni

Contoh:

Ada suatu himpunan $A = \{a, b, c, \}$

Himpunan Bagian dari A adalah

Himpunan Kosong = $\{ \}$

Himpunan 1 Unsur = $\{a\}, \{b\}, \{c\}$

Himpunan 2 Unsur = $\{a, b\}, \{a, c\}, \{b, c\}$

Himpunan 3 Unsur = $\{a, b, c, \}$

Proper subset nya adalah Himpunan 1 Unsur dan Himpunan 2 Unsur

4. Item set adalah Himpunan item atau item-item di I

Contoh:

Ada suatu himpunan $A = \{a, b, c\}$

Item set nya adalah $\{a\}; \{b\}; \{c\}; \{a, b\}; \{a, c\}; \{b, c\}$

5. K- item set adalah Item set yang terdiri dari K buah item yang ada pada I. Intinya K itu adalah jumlah unsur yang terdapat pada suatu Himpunan
- Contoh: 3-item set adalah yang bersifat 3 unsur

6. Item set Frekuensi adalah Jumlah transaksi di I yang mengandung jumlah item set tertentu. Intinya jumlah transaksi yang membeli suatu item set.

Contoh:

- Frekuensi Item set yang sekaligus membeli A dan B adalah 3
- Frekuensi item set yang membeli sekaligus membeli A, B dan C adalah 2

7. Frekuen Item Set adalah item set yang muncul sekurang-kurangnya “sekian” kali di D. Kata “sekian” biasanya di simbolkan dengan Φ . Φ merupakan batas minimum dalam suatu transaksi

Contoh:

Pertama kita tentukan $\Phi = 3$, karena jika tidak di tentukan maka maka frekuen item set tidak dapat di hitung. Jika $\Phi=3$ untuk $\{A, B\}$ apakah frekuen Item set? misalnya Jika kita hitung maka jumlah transaksi yang membeli A sekaligus membeli B adalah 5. Karena $5 \geq 3$ maka $\{A, B\}$ merupakan Frekuen Item set.

8. F_k adalah Himpunan semua frekuen Item Set yang terdiri dari K item.

4. Algoritma – algoritma Asosiasi

a. Algoritma Apriori

Menurut Eko (2012:317) pendekatan dengan algoritma apriori berusaha untuk secara efisien menemukan jumlah *itemset* frekuen. Algoritma ini menjadi pelopor dalam algoritma analisis asosiasi untuk menemukan kandidat *itemset* frekuen dan pembangkitan aturan asosiasi yang dapat dibentuk. Hal utama dalam teorema algoritma apriori menggunakan prinsip: “Jika sebuah *itemset* itu frekuen, semua subset (bagian) dari *itemset* tersebut pasti juga frekuen”.

Algoritma apriori menggunakan pendekatan iterative yang dikenal dengan *level-wise search*, dimana k-kelompok produk digunakan untuk mengeksplorasi (k+1)-kelompok produk atau (k+1) *itemset*.

b. Algoritma Hash Based

Algoritma *hash-based* menggunakan teknik *hashing* untuk menyaring keluar *itemset* yang tidak penting untuk pembangkitan *itemset* selanjutnya. Ketika *support count* untuk kandidat *k-itemset* dihitung dengan menelusuri basis data, algoritma *hash-based* mengumpulkan informasi mengenai (k+1)-*itemset* dengan cara seluruh kemungkinan di-*hash* ke dalam *hash table* dengan menggunakan fungsi *hash*.

c. Algoritma FP-Growth

Algoritma FP-Growth merupakan salah satu alternatif algoritma yang cukup efektif untuk mencari himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data yang besar. Algoritma FP-Growth merupakan algoritma *Association Rules* yang cukup sering dipakai. Algoritma FP-Growth ini dikembangkan dari algoritma apriori. Karakteristik algoritma FP-Growth adalah struktur data yang digunakan adalah tree yang disebut dengan FP-Tree. Dengan menggunakan FP-Tree, algoritma FP-growth dapat langsung mengekstrak *frequent Itemset* dari FP-Tree. Penggalan itemset yang *frequent* dengan menggunakan algoritma FP-Growth akan dilakukan dengan cara membangkitkan struktur data tree atau disebut dengan FP-Tree.

E. Algoritma Apriori

1. Konsep Algoritma Apriori

Algoritma Apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule* (Erwin, 2009). Algoritma Apriori menggunakan pengetahuan frekuensi atribut yang telah diketahui sebelumnya untuk memproses informasi selanjutnya. Pada algoritma Apriori menentukan kandidat yang mungkin muncul dengan cara memperhatikan minimum *support* dan minimum *confidence*. *Support* adalah nilai pengunjung atau persentase kombinasi sebuah *item* dalam *database*.

Sedangkan *confidence* adalah nilai kepastian yaitu kuatnya hubungan antar item dalam sebuah *Apriori*. *Confidence* dapat dicari setelah pola frekuensi munculnya sebuah item ditemukan. Proses utama yang dilakukan dalam algoritma Apriori untuk mendapat *frequent itemset* yaitu (Erwin, 2009) :

1. *Join* (penggabungan)

Proses ini dilakukan dengan cara pengkombinasian item dengan item yang lainnya hingga tidak dapat terbentuk kombinasi lagi.

2. *Prune* (pemangkasan)

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

Prinsip dari Algoritma Apriori antara lain :

- 1) Mengumpulkan item yang tunggal kemudian mencari item yang terbesar.
- 2) Dapatkan *candidate pairs* (2-itemset) kemudian hitung *large pairs* dari masing-masing item.
- 3) Temukan *candidate triplets* (3itemset) dari setiap item dan seterusnya hingga k-itemset.
- 4) Setiap subset dari sebuah *frequent itemset* harus menjadi *frequent*.

Beberapa istilah yang digunakan dalam algoritma apriori antara lain:

- a. *Support* (dukungan): probabilitas pelanggan membeli beberapa produk secara bersamaan dari seluruh transaksi. Support untuk aturan " $X \Rightarrow Y$ " adalah probabilitas atribut atau kumpulan atribut X dan Y yang terjadi bersamaan.
- b. *Confidence* (tingkat kepercayaan): probabilitas kejadian beberapa produk dibeli bersamaan dimana salah satu produk sudah pasti dibeli. Contoh: jika ada n transaksi dimana X dibeli, dan ada m transaksi dimana X dan Y dibeli bersamaan, maka confidence dari aturan if X then Y adalah m/n .
- c. *Minimum support*: parameter yang digunakan sebagai batasan frekuensi kejadian atau support count yang harus dipenuhi suatu kelompok data untuk dapat dijadikan aturan.
- d. *Minimum confidence*: parameter yang mendefinisikan minimum level dari confidence yang harus dipenuhi oleh aturan yang berkualitas.
- e. *Itemset*: kelompok produk.
- f. *Support count*: frekuensi kejadian untuk sebuah kelompok produk atau itemset dari seluruh transaksi.
- g. Kandidat *itemset*: itemset-itemset yang akan dihitung *support count*-nya.
- h. *Large itemset*: itemset yang sering terjadi, atau itemset-itemset yang sudah melewati batas *minimum support* yang telah diberikan.

2. Tahapan Algoritma Apriori

Langkah-langkah algoritma Apriori menurut Fajar (2013:103) :

- Mula-mula $k=1$
- Bangkitkan frequent item set s dengan panjang 1
- Ulangi sampai frequent item set baru yang dikenali
 - Bangkitkan kandidat item set dengan panjang $(k+1)$ dari frequent item set s dengan panjang k .
 - Pangkas kandidat itemset s yang berisi subset dengan panjang k yang infrequent.
 - Hitung support dari tiap kandidat dengan memindai *database*.
 - Eliminasi kandidat yang infrequent, hanya menyisakan yang frequent.

Sejalan dengan pendapat Fajar (2013:103) di atas Eko (2012:121) mengemukakan algoritma Apriori dalam bentuk pseudokode. Jika C_k menyatakan kandidat k -itemset, dan F_k menyatakan set k -itemset yang frequent:

1. $k=1$
2. $F_k = \{i | i \in I \wedge \sigma(\{i\}) \geq N \times \text{minsup}\}$ %mencari semua 1-itemset yang frequent
3. Ulangi langkah 4 sampai 10, sampai $F_k = \emptyset$
4. $k = k + 1$
5. $C_k = \text{apriori_gen}(F_k)$ %bangkitkan kandidat itemset
6. Untuk setiap transaksi $t \in T$, lakukan langkah 7 sampai 9
7. $C_t = \text{subset}(C_k, t)$ %identifikasi semua kandidat yang dimiliki t
8. Untuk semua kandidat itemset $c \in C_t$, lakukan langkah 9
9. $\sigma(c) = \sigma(c) + 1$ %naikkan support count
10. $F_k = \{c | c \in C_k \wedge \sigma(c) \geq N \times \text{minsup}\}$ %ekstrak k -itemset yang frequent

1. Awalnya algoritma membaca semua transaksi satu kali untuk menentukan support setiap item. Selanjutnya pada langkah 2 akan dipilih 1-*itemset* yang frekuen (langkah 1 dan 2).
2. Selanjutnya secara iterative algoritma akan membangkitkan kandidat k-*itemset* yang baru menggunakan (k-1)-*itemset* yang frekuen yang ditemukan pada iterasi sebelumnya (langkah 5)
3. Untuk menghitung *support count* dari setiap kandidat, algoritma perlu melakukan pembacaan tambahan pada set data (langkah 6 sampai 9). Fungsi subset() digunakan untuk menentukan semua kandidat itemset dalam C_k yang mengisi setiap transaksi t .
4. Setelah menghitung support, algoritma akan membuang semua kandidat itemset yang nilai *support count*-nya kurang dari minsup (langkah 10).
5. Algoritma berhenti ketika tidak ada lagi itemset yang bisa dibangkitkan, misalnya $F_k = \emptyset$

Berdasarkan jurnal Informatika Marsela dan Veronika (2004) tahapan yang dilakukan algoritma apriori untuk membangkitkan *large itemset* adalah sebagai berikut:

1. Menelusuri seluruh *record* di basis data transaksi dan menghitung *support count* dari tiap item. Ini adalah kandidat 1-*itemset*, C_1 .
2. Large 1-*itemset* L_1 dibangun dengan menyaring C_1 dengan *support count* yang lebih besar sama dengan *minimum support* untuk dimasukkan kedalam L_1 .

3. Untuk membangun L2, algoritma apriori menggunakan proses join untuk menghasilkan C2.
4. Dari C2, 2-itemset yang memiliki *support count* yang lebih besar sama dengan *minimum support* akan disimpan ke dalam L2.
5. Proses ini diulang sampai tidak ada lagi kemungkinan k-itemset.

F. Hubungan Market Basket Analysis dengan Asosiasi pada Algoritma Apriori

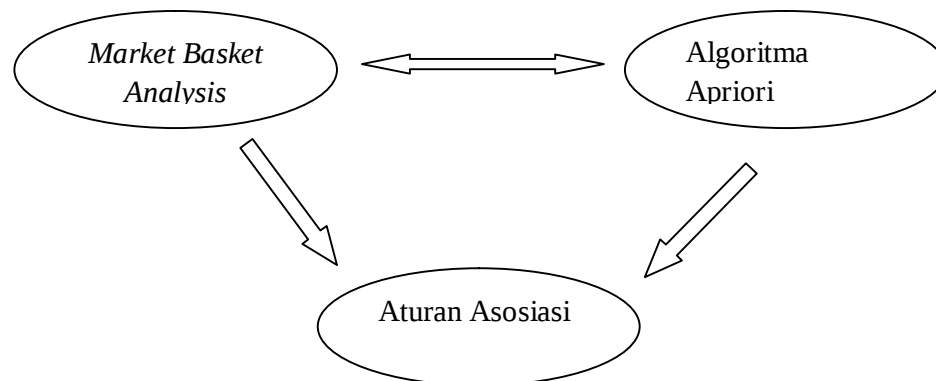
Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. Di dalam bidang usaha retail metode *association rule mining* ini lebih dikenal dengan istilah analisa keranjang belanja (*market basket analysis*).

Market Basket Analysis adalah salah satu teknik pemodelan dalam *data mining* berdasarkan teori yang mana jika anda membeli suatu grup *item*, anda akan memiliki kemungkinan membeli *itemset* yang lain (*data mining concept and technique*, Jiawei Han).

Algoritma Apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule* (Erwin, 2009).

Dari definisi diatas dapat disimpulkan bahwa algoritma apriori merupakan salah satu proses di dalam *Market Basket Analysis*. Hasil atau output dari *Market Basket Analysis* dengan algoritma apriori berupa pola keterkaitan antar item atau yang disebut dengan aturan asosiasi.

Hubungan antara Market Basket Analisis dengan Asosiasi pada Algoritma Apriori dapat dilihat pada Gambar 3 berikut:



Gambar 3. Hubungan MBA dengan Asosiasi dan Algoritma Apriori

G. Software Aplikasi data mining untuk Pengujian

1. WEKA 3.7.4 Data Mining Tool

Feri dan Dominikus (2010: 63) menyatakan bahwa weka merupakan aplikasi *data mining* yang berbasis *open source* (GPL) dan berengine java.

Aplikasi ini dikembangkan pertama kali oleh sebuah Universitas di Selandia Baru yang bernama Universitas Waikato sebelum menjadi bagian di Pentaho (baca akusisi). Software yang mulai dikembangkan sejak tahun 1994 telah menjadi software *data mining open source* yang paling populer. Banyaknya algoritma *data mining* dan *machine learning*, kemudahan dalam penggunaan, ditambah lagi selalu *up-to-date* dengan algoritma-algoritma baru yang muncul menjadikan software ini banyak digunakan.

Weka mendukung beberapa format file untuk input-nya, diantaranya:

- a. *Comma Separated Values* (CSV), merupakan file teks dengan pemisah tanda koma (,) yang cukup umum digunakan. Membuat

filenya dengan menggunakan Microsoft Excel atau dengan menggunakan notepad.

- b. Format C45, merupakan format file yang bisa diakses dengan menggunakan aplikasi Weka.
- c. *Attribute-Relation File Format (ARFF)*, merupakan tipe file teks yang berisi berbagai *instance* data yang berhubungan dengan suatu set atribut data yang dideskripsikan juga dalam file tersebut.

2. Tanagra Data Mining Tool

Tanagra merupakan salah satu software data mining yang didalamnya disediakan beberapa metoda data mining mulai dari mengeksplorasi analisis data, pembelajaran statistik, pembelajaran mesin, dan database.

Tidak seperti software data mining kebanyakan, tanagra merupakan suatu software berbasis open source di mana semua orang dapat mengakses source codenya, dan menambahkan algoritma mereka sendiri, sejauh dia setuju dan menyesuaikan dengan lisensi pendistribusian softwarenya.

a. Data Input dalam Software Tanagra

1) Microsoft Excel

Kebanyakan data yang diolah oleh Tanagra bersumber dari data yang di tulis dari Microsoft Excel.

2) TXT

Data yang di import menggunakan format TXT berupa data yang di buat dengan menggunakan Excel kemudian di exspor nya ke txt.

3) ARFF

Data yang formatnya Arff (Attribute-Relation File Format) ini merupakan format yang digunakan oleh Weka dan Tanagra juga bisa menggunakan langsung.

b. Data Output dalam Software Tanagra

1) Binary description of the stream diagram(*.bdm).

File yang berformat (*.bdm) ini hanya dapat di manfaatkan oleh Tanagra. Keuntungan utama dari format ini adalah bahwa data yang di impornya sekali dan hanya sekali. Di sisi lain, ketidaknyamanan utama format ini adalah bahwa penyusunan analisis diagram yang definitif yang di definisikan atas data yang diimpor. Jadi jika data berubah, dengan menambahkan beberapa catatan misalnya, data harus diimpor lagi, maka diagram harus didefinisikan ulang. Jadi kesimpulan format ini adalah data yang di hasilkan tidak seharusnya ada perubahan lagi.

2) Textual description of the stream diagram (*.TDM).

Format ini, berdasarkan format file INI Windows, menjelaskan dalam sebuah file teks analisis penyusunan diagram. Jadi file ini dapat dibuka dan diperiksa dengan editor teks apapun.

Keuntungan format ini yaitu hanya ada referensi ke data dalam file yang disimpan, jika data nya terjadi perubahan, maka eksekusi berikutnya akan bekerja pada versi baru dari data, dan menghasilkan hasil yang diperbarui, file yang dihasilkan menghormati spesifikasi INI, sehingga memungkinkan untuk mendefinisikan diagram baru, tanpa membuka Tanagra. Kekurangan utama dari format penyimpanannya adalah jika kebutuhan untuk mengimpor data setiap kali maka Anda menjalankan diagram aliran.

H. Jurnal Terkait

Berdasarkan penulisan yang dilakukan oleh Goldie Gunadi dan Dana Indra Sensuse pada tahun 2012 dalam jurnalnya yang berjudul *penerapan metode data mining market basket analysis terhadap data penjualan produk buku dengan menggunakan algoritma apriori dan frequent pattern growth (fp-growth)*, dikemukakan bahwa penulisan tersebut berisi tentang perbandingan algoritma apriori dengan *frequent pattern growth (fp-growth)*. Dalam penelitian ini disebutkan tingkat kekuatan aturan-aturan asosiasi menunjukkan bahwa aturan-aturan asosiasi yang dihasilkan oleh algoritma Apriori memiliki tingkat kekuatan yang lebih tinggi dibandingkan dengan yang dihasilkan oleh algoritma FP-growth. Tingkat kekuatan dari aturan asosiasi ditentukan oleh nilai support yang mewakili aspek generalitas dan nilai confidence yang mewakili aspek reliabilitas.

Penulisan lain yang dilakukan oleh Muhammad Ikhsan, ST., M.Kom dkk pada tahun 2000 dalam jurnalnya yang berjudul *Penerapan Associaton Rule dengan Algoritma Apriori Pada Proses Pengelompokan Barang di Perusahaan Retail* menjelaskan penggunaan data mining dengan teknik *Association rule* untuk menggali pola hubungan atribut-atribut dan *frequent itemset* dalam database Retail. Paradigma apriori digunakan untuk mencari large itemset dalam penetapan *association rule*. Dalam penelitian ini disimpulkan integrasi *association rule* dengan paradigma apriori telah berhasil menemukan sejumlah pola hubungan antar atribut dalam database retail.

Penulisan yang dilakukan oleh Dewi Kartika Pane pada tahun 2013 dalam jurnalnya yang berjudul *Implementasi Data Mining pada Penjualan Produk Elektronik dengan Algoritma Apriori (Studi Kasus: KreditPlus)*. Dalam penulisan ini penulis mengaplikasikan algoritma apriori untuk mengetahui penjualan produk elektronik yang banyak terjual dan mengetahui sejauh mana algoritma apriori dapat membantu pengembangan strategi pemasaran. Data yang digunakan dalam penelitian ini adalah data penjualan *KreditPlus* selama 1 tahun terakhir, dataset diperoleh berdasarkan penjualan bulanan yang diambil dari 3 penjualan teratas setiap bulannya. Penulisaan ini menjelaskan bagaimana mengimplementasikan algoritma apriori pada data penjualan barang, dalam penelitian ini disimpulkan algoritma apriori dapat membantu mengembangkan strategi pemasaran.

Penulisan yang dilakukan oleh Triana Faradila Putri, M. Arif Rahman, dan Edy Santoso pada tahun 2010 dalam jurnalnya dengan judul *Identifikasi*

Pengkategorian Mushroom Menggunakan Algoritma Apriori dengan Algoritma Perfect Hashing and Pruning (php). Penelitian ini bertujuan untuk membuat suatu sistem untuk pengkategorian *mushroom* dengan menemukan association rule sesuai dengan nilai minimum *confidence* dan nilai minimum *support*. Penelitian ini menyimpulkan semakin besar nilai minimum *support* dan nilai minimum *confidence* maka semakin sedikit jumlah *rule* yang ditemukan dan sebaliknya, semakin kecil nilai minimum *support* dan nilai minimum *confidence* maka *rule* yang dihasilkan lebih banyak dan beragam.

I. Teori tentang Manajemen Pemasaran dan Penjualan

Pemasaran merupakan suatu disiplin ilmu yang dipergunakan oleh perusahaan untuk memenuhi kebutuhan konsumen menjadi peluang yang menghasilkan laba perusahaan. Definisi pemasaran menurut Kotler (2003:9), adalah proses sosial yang didalamnya individu dan kelompok mendapatkan apa yang mereka butuhkan dan inginkan dengan menciptakan, menawarkan dan secara bebas mempertukarkan produk yang bernilai dengan pihak lain.

Manajemen pemasaran Kotler (2003:9) adalah seni dan ilmu memilih pasar sasaran dan mendapatkan, menjaga serta menumbuhkan pelanggan dengan menciptakan, menyerahkan dan mengkomunikasikan nilai pelanggann yang unggul.

Dengan demikian, dapat diketahui bahwa tugas manajemen pemasaran bukan hanya menawarkan barang atau jasa yang sesuai dengan kebutuhan dan keinginan pasarnya, menetapkan harga yang efektif, komunikasi dan distribusi untuk memberikan informasi, mempengaruhi dan melayani pasarnya tetapi

lebih dari itu. Tugas manajemen pemasaran adalah mempengaruhi tingkat, waktu dan komposisi permintaan untuk membantu perusahaan mencapai sasarannya.

Penjualan merupakan bagian dari manajemen pemasaran. Penjualan adalah suatu usaha yang terpadu untuk mengembangkan rencana-rencana strategis yang diarahkan pada usaha pemuasan kebutuhan dan keinginan pembeli, guna mendapatkan penjualan yang menghasilkan laba (Marwan, 1991).

Menurut Hary (2003) dalam artikelnya yang berjudul *Belajar Manajemen kategori* menyatakan manajemen penjualan yang baik dapat melaksanakan beberapa kegiatan atau pekerjaan yang terdiri dari :

1. Merencanakan pembelian, caranya :
 - a. menganalisa konsumen, toko/*account* pesaing, dan lingkungan sekitar untuk mengetahui siapa dan seperti apa target konsumennya
 - b. *setting* strategi, kebijakan, dan *goal* yang diinginkan untuk memuaskan target konsumennya
 - c. *setting assortment* (pemilihan produk) dan seberapa lebar *range* kategorinya
 - d. menentukan konsep *merchandising* dan spesifikasinya seperti apa
 - e. menentukan dari tiap produk dalam kategorinya, akan disediakan (dibeli) berapa banyak?

2. Membeli barang

Dalam kegiatan membeli, maka hal yang harus diperhatikan oleh pihak manajemen adalah :

1. memilih dan mengelola *supplier* (pemasok)
2. tahu trend di pasar seperti apa, dalam memilih produk-produk apa saja yang laku, punya *turn over* tinggi, dan layak jual.
3. negosiasi *trading term* (kontrak kerjasama jual-beli antara *supplier-retailer* selama 1 tahun), harga, dan kondisi-kondisi lainnya.
4. menambah atau mengurangi *range* dan jumlah produk dari kategori yang dikelola, dan berhak menghapus produk yang tidak laku.

3. Distribusi

Pekerjaan yang dilakukan dalam manajemen distribusi adalah :

- a. menganalisa dan menentukan strategi distribusi dan *warehousing* (penyimpanan) dengan tepat.
- b. menentukan *delivery* ke toko tepat waktu dan jumlah (khusus *supplier* dengan *warehouse buying*).
- c. mengontrol *service level* dari *supplier* yang dikelola.
- d. monitor secara berkala mekanisme FIFO (*first in first out* – barang pertama masuk akan keluar lebih dahulu) di gudang dan took
- e. melakukan juga *quality control* atas produk yang dijual, dan prosedur retur produk.

4. Penataan produk di toko / supermarket

- a. memastikan lokasi penempatan produk dan tata letaknya di toko

- b. evaluasi tata letak dan *display*, segmen atau sub kategori apa yang ditata berdekatan dengan kategori mana, dsb.
 - c. membuat perencanaan promosi dan kegiatan berbau *marketing* baik secara internal maupun dengan *supplier*.
 - d. memastikan toko tahu cara *handling produk* dengan benar
5. Control terhadap pengelolaan produk yang dijual
- a. memonitor dan analisa pasar, penjualan, dan stock.
 - b. evaluasi CSL (*customer service level*)
 - c. mendapatkan *feed back* dari toko.
 - d. mengelola *stock level* dan mekanisme pengorderan kembali (*stock refill*).
 - e. perencanaan kembali jika ada perubahan, untuk tetap mencapai tujuan utama.

Supermarket Robinson berada dibawah naungan PT. Ramayana Lestari Sentosa. Perusahaan ini dirintis oleh suami istri Paulus Tumewu di Ujung Pandang. Maka pada tahun 1978 didirikanlah outlet perintis "Ramayana Fashion Store" di Jl. H.Agus Salim, Jakarta Pusat yang dikenal dengan " R 1 ". Dari sinilah bisnis eceran Ramayana dan Robinson Group menggurita semakin aktif mendekati konsumen dan memiliki banyak cabang di berbagai daerah di Indonesia, salah satunya adalah Supermarket Robinson Plaza Andalas.

Sejak bisnis dimulai pada tahun 1978, Ramayana bergerak lambat tapi tumbuh dengan mantap. Meskipun krisis nampaknya belum berakhir, namun pertumbuhan bisnis yang cepat dapat melewati masa pra krisis lebih cepat

dibandingkan kompetitor terdekat dan dianugerahi penghargaan sebagai hasilnya.

Ramayana mempunyai 3 prinsip yaitu :

1. Selalu menyediakan produk berkualitas tinggi dengan harga terjangkau.
2. Selalu mengutamakan pelayanan yang terbaik kepada pelanggan dan membina hubungan baik dengan mitra usaha.
3. Bersama-sama meningkatkan kualitas sumber daya manusia sehingga Ramayana dapat berpartisipasi dalam pembangunan negara dan menciptakan kesejahteraan seluruh karyawan.

Supermarket Robinson Plaza Andalas melakukan berbagai cara untuk meningkatkan daya saingnya dibanding supermarket lain, salah satunya yaitu dengan meningkatkan kenyamanan pelanggan atau konsumen untuk menemukan produk yang dicari dengan mudah yaitu dengan cara :

1. Menata rak sedemikian rupa agar konsumen bisa dengan segera menemukan produk yang dicari
2. Membagi produk berdasarkan kategori dan sub kategori untuk ditata di rak-rak atau divisi.
3. Membuat nama atau *signage* besar di rak untuk membantu menemukan lorong rak yang dicari dengan segera.

Tata letak barang di Supermarket Robinson diatur dengan pengelompokan produk. Produk-produk yang disediakan akan dibagi menjadi dua bagian, *food* dan *non food*. Setelah pintu masuk pelanggan akan menemui produk-produk *non-food*, berlanjut menemui produk *food*, dan produk *fresh*

dibagian paling ujung. Produk *fresh* diletakkan diujung, untuk menjaga kesegaran-nya. Misalnya : produk-produk beku.

Selain itu usaha yang dilakukan pihak manajemen supermarket untuk meningkatkan penjualan yaitu dengan memberikan penawaran harga yang lebih murah saat masa promosi.

BAB V

PENUTUP

A. Simpulan

Berdasarkan penelitian yang dilakukan maka dapat di dapat kesimpulan sebagai berikut :

1. Penerapan metode analisis asosiasi pada data penjualan barang dengan algoritma apriori telah menghasilkan 130 aturan asosiasi yang mana diambil 5 aturan asosiasi yang tergolong *strong rules* yaitu aturan yang memiliki nilai *support* dan *confidence* paling tinggi.
2. Pola kombinasi item yang dihasilkan yang berupa aturan asosiasi dapat membantu pihak manajemen Supermarket Robinson Plaza Andalas untuk menemukan keterkaitan atau pola kemunculan barang dalam transaksi penjualan, yang pada akhirnya dapat digunakan untuk menyusun strategi penjualan.
3. Analisis asosiasi dengan *tool* data mining WEKA 3.7.4 menghasilkan 130 aturan asosiasi yang mana aturan yang dihasilkan diurutkan berdasarkan nilai *support* dan *confidence* tertinggi sehingga didapat 5 aturan yang termasuk kelompok *Strong Rules* dan dapat diketahui pola penjualan selama satu tahun terakhir di Supermarket Robinson Plaza Andalas.
4. Analisis asosiasi dengan menggunakan *tool* data mining Tanagra 1.4 menghasilkan 130 aturan asosiasi yang diurutkan berdasarkan nilai LIFT. Aturan asosiasi yang dihasilkan dengan Tanagra 1.4 belum bisa

menemukan *Strong Rules* karena aturan tidak diurutkan berdasarkan nilai *confidence*.

5. Analisis dengan menggunakan WEKA lebih akurat dibandingkan dengan menggunakan Tanagra karena dengan WEKA kita bisa langsung menemukan *Strong Rules* pada aturan asosiasi yang dihasilkan sedangkan, dengan Tanagra kita harus mengeksplor data terlebih dahulu untuk mendapatkan *Strong Rules*.

B. Saran

Berdasarkan hasil penelitian yang telah dilakukan maka saran yang dapat diberikan :

1. Pihak manajemen Supermarket Robinson Plaza Andalas bisa menjadikan pola aturan asosiasi yang dihasilkan sebagai acuan dalam melakukan usaha-usaha peningkatan penjualan.
2. Diharapkan penelitian selanjutnya menggunakan data yang lebih besar lagi sehingga *rules* yang dihasilkan lebih beragam dan dapat menemukan pola yang lebih baik.

DAFTAR PUSTAKA

- Budhi Santosa dkk. 2007. *Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis*, GRAHA ILMU.
- Dedy Suryadi dan Sani Susanto. 2010. *Pengantar Data Mining menggali pengetahuan dari bongkahan data*, Yogyakarta: ANDI.
- Dewi Kartika Pane. 2013. "Implementasi Data Mining pada Penjualan Produk Elektronik dengan Algoritma Apriori (Studi kasus: KreditPlus)". *Pelita Informatika Budi Darma* Vol. 4 No. 3, Agustus 2013.
- Eko Prasetyo. 2012. *Data mining Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: ANDI.
- Erwin, 2009. "Analisis Market Basket Dengan Algoritma Apriori dan FP Growth". *Jurnal Generik* Vol. 4, Juli 2009.
- Fajar Astuti. 2013. *DATA MINING*. Yogyakarta: ANDI.
- Feri Sulianta dan Dominikus Juju. 2010. *Data Mining Meramalkan Bisnis Perusahaan*. Jakarta: PT Elex Media Komputindo.
- Goldie Gunadi dan Dana Indra. 2012. "penerapan metode data mining market basket analysis terhadap data penjualan produk buku dengan menggunakan algoritma apriori dan frequent pattern growth (fp-growth)". *Jurnal TELEMATIKA MKOM* Vol.4 No.1, Maret 2012.
- Han, Jiawei dan Micheline Kamber. 2006. *Data Mining Concepts and Techniques 2nd Edition*. San Fransisco : Morgan Kaufmann publisher.
- Herlawati, Prabowo P. J dan Rahmadya T. H. 2013. *Penerapan Data Mining dengan MATLAB*. Bandung : Rekayaa Sains.
- Johan Oscar Ong. 2013. Implementasi Algoritma K-Means Clustering untuk Menentukan Strategi Marketing President University. *Jurnal*. Bekasi: Program Studi Teknik Industri, President University.
- Larose, Daniel T. 2005. *Discovery Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc. New Jersey.

- Muhammad Ikhsan, ST., M.Kom dkk. 2000. “Penerapan Associaton Rule dengan Algoritma Apriori Pada Proses Pengelompokan Barang di Perusahaan Retail”.
- Sugiyono. 2009. *Metode Penelitian Pendidikan*. Bandung:Alfabeta.
- Triana Fadila Putri. 2010. “Identifikasi Pengkategorian Mushroom Menggunakan Algoritma Apriori dengan Algoritma Perfect Hashing and Pruning (php)”.
- Wu, Xindong & Vipin Kumar. 2009. *The Top Ten Algoritms in Data Mining*. New York. CRC Press.
- Yusuf yogi & Rian Pratikto, Gerry. 2006. “Penerapan Data Mining Dalam Penentuan Aturan Asosiasi Antar Jenis Item”. Jurnal Teknologi Informasi, 17 Juni 2006.