

Analisa Data Mining Menggunakan *Market Basket Analysis* untuk Mengetahui Pola Beli Konsumen

Alkadri Masnur

Teknik Informatika, STMIK Amik Riau
alkadrimasnur@stmik-amik-riau.ac.id

Abstrak

Pada saat sekarang ini data tidak dapat dipisahkan dari kehidupan sehari-hari dan merupakan salah satu sumber daya yang sangat berharga. Data - data yang terlibat dalam setiap transaksi penjualan pada toko bangunan sangat banyak, sehingga terjadilah tumpukan data yang dibiarkan saja. Untuk itu digunakanlah metode market basket analysis untuk menyelesaikan masalah yang dihadapi yang nanti menghasilkan rule-rule yang sangat berguna untuk memberi informasi kepada toko bangunan tentang barang-barang yang laris terjual serta barang-barang yang jarang dibutuhkan oleh konsumen. Hasil akhir yang diperoleh dari penelitian ini nanti adalah pihak toko bangunan dapat mengetahui pola beli konsumen dan juga mengetahui barang yang laris dijual. Selain itu juga untuk meningkatkan pelayanan dan meningkatkan penjualan.

Kata Kunci : data mining, market basket analysis, pola beli konsumen, toko bangunan

1. Pendahuluan

Data tidak dapat dipisahkan dari kehidupan sehari-hari dan merupakan salah satu sumber daya yang sangat berharga. Suatu organisasi baik besar maupun kecil dapat dibanjiri dengan berbagai macam data, tidak terkecuali dengan Toko Bangunan. Data - data yang terlibat dalam setiap transaksi penjualan pada toko ini, seperti data item yang dibeli, waktu pembelian, jumlah item yang dibeli, harga item, dan data konsumen. Karena banyaknya data dalam setiap transaksi dan pertumbuhan data yang pesat sehingga tidak jarang kumpulan data yang dibiarkan begitu saja. Padahal kita bisa menambang informasi-informasi dari kumpulan data itu dan menjadikannya informasi yang berguna untuk mendukung keputusan ataupun membantu dalam menentukan strategi pemasaran.

Berlatar belakang pada permasalahan diatas, penulis melakukan analisa data dengan menggunakan teknik *Market Basket Analysis* atau sering juga disebut

Association rule. Dengan adanya teknik ini kita dapat mengidentifikasi hubungan antara item dan kita juga dapat menentukan item mana saja yang paling sering dibeli sekaligus item mana saja yang sering dibeli secara bersamaan. Setelah proses identifikasi maka kita dapat mengetahui pola beli konsumen.

Pola-pola atau rules yang diperoleh nantinya dapat dijadikan acuan dalam menentukan penyusunan letak item secara rapi menurut kombinasi barang yang paling sering dibeli dan saling berhubungan, dengan tujuan meningkatkan pelayanan dan penjualan barang dengan strategi pemasaran yang tepat.

2. Landasan Teori

2.1 *Knowledge Discovery in Database (KDD)*

Knowledge Discovery and Data Mining (KDD) adalah proses yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan yang berguna. *Data Mining tools* memperkirakan perilaku dan tren masa depan, memungkinkan bisnis untuk membuat keputusan yang proaktif dan berdasarkan pengetahuan. *Data Mining tools* mampu menjawab permasalahan bisnis yang secara tradisional terlalu lama untuk diselesaikan. *Data Mining tools* menjelajah database untuk mencari pola tersembunyi, menemukan informasi yang prediktif yang mungkin dilewatkan para pakar karena berada di luar ekspektasi mereka [1].

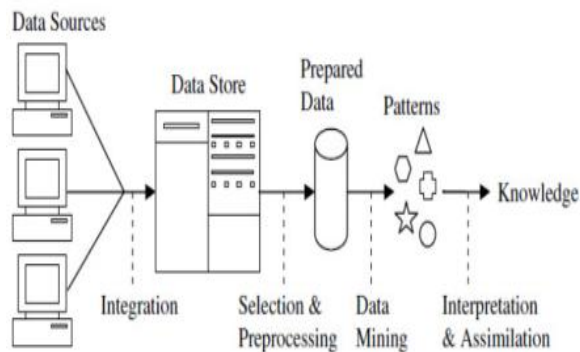
Proses dalam KDD adalah proses yang digambarkan pada dan terdiri dari rangkaian proses iteratif sebagai berikut [1]:

1. Data cleaning adalah menghilangkan noise dan data yang inkonsisten.
2. Data integration adalah menggabungkan data dari berbagai sumber data yang berbeda.
3. Data selection adalah mengambil data yang relevan dengan tugas analisis dari database.
4. Data transformation adalah mentransformasi atau menggabungkan data ke dalam bentuk yang sesuai untuk penggalian lewat operasi summary atau aggregation.

5. *Data mining* adalah proses esensial untuk mengekstrak pola dari data dengan metode cerdas.
6. *Pattern evaluation* adalah mengidentifikasi pola yang menarik dan merepresentasikan pengetahuan berdasarkan interestingness measures.
7. *Knowledge presentation* adalah penyajian pengetahuan yang digali kepada pengguna dengan menggunakan visualisasi dan teknik representasi pengetahuan.

2.2 Data Mining

Data mining adalah proses menganalisis data dari perspektif yang berbeda dan menjadi sebuah informasi yang berguna yang dapat digunakan untuk meningkatkan keuntungan [6] *Data mining* mengacu pada analisis dari jumlah besar data yang disimpan dalam computer. Data pertambangan telah disebut analisis data eksplorasi, antara lain. Misal data yang dihasilkan dari cash register, dari pemindaian, dari database tertentu topik di seluruh perusahaan, dieksplorasi, dianalisis, dikurangi, dan digunakan kembali [2].



Gambar 1. Proses data mining

(a) Pengelompokan Data Mining

Data Mining dapat dikelompokkan ke dalam beberapa kategori berdasarkan tugas dan fungsinya, yaitu *Description*, *Estimation*, *Prediction*, *Classification*, *Clustering*, dan *Association* [8].

1 Deskripsi (*Description*)

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau Kecenderungan.

2 Estimasi (*Estimation*)

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

prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi (*Prediction*)

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi (*Classification*)

Suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Aturan-aturan tersebut digunakan pada data-data baru untuk diklasifikasi. Teknik ini menggunakan *supervised induction*, yang memanfaatkan kumpulan pengujian dari *record* yang terklasifikasi untuk menentukan kelas-kelas tambahan.

5. Pengklasteran (*Clustering*)

Clustering adalah teknik KDD penting dengan berbagai aplikasi, seperti pemasaran dan segmentasi pelanggan. *Clustering* biasanya kelompok data ke set sedemikian rupa bahwa *intra-cluster* kesamaan dimaksimalkan dan sementara antar-klaster kesamaan diminimalkan *Clustering* adalah belajar tanpa pengawasan. Algoritma *Clustering* memeriksa data untuk menemukan kelompok - kelompok barang yang mirip. Misalnya, perusahaan asuransi mungkin pelanggan kelompok sesuai dengan pendapatan, usia, jenis kebijakan yang dibeli, pengalaman klaim sebelumnya dalam aplikasi diagnosis kesalahan, kesalahan listrik mungkin dikelompokkan menurut nilai-nilai variabel kunci tertentu [2].

6. Association

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 tesis ini. Contoh dari aturan asosiatif dari analisa pembelian di suatu pasar swalayan adalah bisa diketahui berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

(b) Mining Association Rule

Association rule mining adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu

kombinasi item. Aturan Asosiasi bertujuan untuk menemukan barang-barang yang sering dibeli yang terlihat dalam database. Sebagai contoh seorang manajer toko mungkin ingin tahu seberapa besar kemungkinan pelanggan akan membeli bola tenis jika ia membeli raket tenis. Pengetahuan tersebut dapat membantu manajer toko dalam pengaturan produk, perencanaan ruang rak dan pelaksanaan yang efektif dari strategi promosi produk.

Tulisan ini merupakan masing-masing item dengan salah satu dari dua variabel *Boolean* yaitu 0 dan 1, di mana 0 mewakili ketidakhadiran 1 mewakili kehadiran. Jadi setiap item dapat direpresentasikan dalam keranjang oleh vektor *Boolean*. Vektor *Boolean* kemudian dianalisis untuk membeli pola yang mencerminkan item yang sering terkait satu sama lain. Pola-pola ini dapat direpresentasikan dalam bentuk aturan asosiasi [6].

2.3 Market Basket Analysis (Association Rule)

Association Rule sering juga disebut dengan *Market Basket Analysis* karena kegunaannya dalam pembuatan model perilaku pembelian oleh konsumen, dan menganalisis perilaku konsumen. *Association Rule* adalah suatu prosedur dalam analisis keranjang pasar (*Market Basket Analysis*) untuk menemukan aturan asosiasi yang memenuhi syarat minimum untuk *support* dan *confidence* (minimum *support* dan minimum *confidence*) [9]. Analisis asosiasi atau *association rule mining* adalah teknik *Data Mining* untuk menemukan aturan asosiatif antara suatu kombinasi item. Algoritma aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian *Data Mining*, untuk menghasilkan pengetahuan.

Association Rule merupakan salah satu metode yang bertujuan mencari pola yang sering muncul di antara banyak transaksi, dimana setiap transaksi terdiri dari beberapa item sehingga metode ini akan mendukung sistem rekomendasi melalui penemuan pola antar item dalam transaksi-transaksi yang terjadi metodologi dasar analisis asosiasi terbagi menjadi dua tahap [3]:

1. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam database. Nilai *support* sebuah item diperoleh dengan rumus berikut:

$$\text{Support}(A) = \frac{\text{Jumlah Transaksi Mengandung } A}{\text{Total Transaksi}} \quad (1)$$

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

$$\text{Support}(A,B) = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Total Transaksi}} \quad (2)$$

2. Pembentukan aturan asosiatif

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A_B Nilai *confidence* dari aturan A_B diperoleh dari rumus berikut:

$$\text{Confidence } P(B|A) = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Jumlah Transaksi Mengandung } A} \quad (3)$$

2.4 Algoritma Apriori

Algoritma apriori adalah algoritma analisis keranjang pasar yang digunakan untuk menghasilkan aturan asosiasi, dengan pola "if-then". *Algoritma apriori* menggunakan pendekatan iteratif yang dikenal dengan level-wise search, dimana k-kelompok produk digunakan untuk mengeksplorasi (k+1)-kelompok produk atau (k+1)-itemset [1].

Beberapa istilah yang digunakan dalam *algoritma apriori* antara lain [1]:

- Support* (dukungan): *probabilitas* pelanggan membeli beberapa produk secara bersamaan dari seluruh transaksi. *Support* untuk aturan "X=>Y" adalah *probabilitas* atribut atau kumpulan atribut X dan Y yang terjadi bersamaan.
- Confidence* (tingkat kepercayaan) yaitu *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.
- Minimum support* yaitu parameter yang digunakan sebagai batasan frekuensi kejadian atau *support count* yang harus dipenuhi suatu kelompok data untuk dapat dijadikan aturan.
- Minimum confidence* yaitu parameter yang mendefinisikan minimum level dari *confidence* yang harus dipenuhi oleh aturan yang berkualitas.
- Itemset* yaitu kelompok produk.
- Support count* yaitu frekuensi kejadian untuk sebuah kelompok produk atau *itemset* dari seluruh transaksi.
- Kandidat *itemset* yaitu *itemset-itemset* yang akan dihitung *support count*-nya.
- Large itemset* yaitu *itemset* yang sering terjadi, atau *itemset-itemset* yang sudah melewati batas *minimum support* yang telah diberikan.

4. Analisa dan Perancangan

4.1. Analisa

Analisa sistem mendefinisikan bagaimana memahami dan menspesifikasi dengan detail apa yang harus dilakukan oleh sistem. Sedangkan sistem desain

diartikan sebagai menjelaskan dengan detail bagaimana bagian-bagian dari sistem informasi diimplementasikan. Sehingga Analisa dan desain sistem informasi (ANSI) bisa didefinisikan sebagai Proses organisasional kompleks dimana sistem informasi berbasis komputer diimplementasikan.

Dalam melihat pola permintaan konsumen yang ada di Toko Bangunan Gunung Saiyo menggunakan banyak metode. Metode yang sering digunakan dalam melihat pola permintaan konsumen yaitu dari laporan transaksi produk.

Pola permintaan konsumen dari laporan transaksi pembelian dilihat dari laporan pembelian produk, dari laporan tersebut pemilik menganalisa dengan membandingkan dengan laporan penjualan setiap periode barang yang akan dilihat pola permintaannya, dari laporan penjualan setiap periode itulah pemilik dapat melihat pola permintaan dari produk yang diinginkan. Untuk menentukan produk yang dipilih untuk dilihat polanya, pemilik hanya dapat memilih satu atau beberapa produk tanpa ada keterkaitan antar produk yang dipilih tersebut.

Model yang digunakan dalam penyelesaian ini adalah model *Market Basket Analysis*. Model *Market Basket Analysis* merupakan model yang pertama yang menyaring produk dengan tingkat asosiasi antar produk. Produk dihasilkan dari *market basket* ini tidak hanya menghasilkan tingkat permintaan produk yang berhubungan.

Proses analisis pola permintaan konsumen diawali dengan pengambilan data transaksional, data transaksi yang di ambil pada periode tertentu diolah dengan data mining menggunakan metode *Market Basket* dan untuk mengetahui asosiasi berupa tingkat *support* dan *confidence* produk.

4.2. Perancangan

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

4.2.1. Pengumpulan Data

Dalam tahap ini, data di dapatkan dari hasil wawancara penulis dengan pihak toko bangunan Gunung Saiyo. Dari hasil wawancara tersebut di dapatlah data transaksi penjualan dalam bentuk file *Microsoft Excel* sebagaimana ditampilkan pada tabel 1.

Tabel 1. Data transaksi penjualan toko bangunan Gunung Saiyo

Transaksi	Item yang di beli
1	Seng, Semen, Paku, Pasir, Batu Bata
2	Triplek, Cat
3	Batu Bata, Besi, Semen, Cat, Triplek, Kuas, Pasir
4	Semen, Besi, Pasir, Cat
5	Seng, Paku
6	Semen, Batu Bata
7	Pasir, Semen, Besi
8	Besi, Triplek, Kawat, Paku
9	Kayu, Paku, Cat, Kuas
10	Pasir, Semen, Batu Bata
11	Semen, Pasir, Seng, Pvc, Paku
12	Kuas, Thiner
13	Batu Bata, Semen, Besi, Kawat
14	Seng, Paku, Semen, Pasir, Triplek, Kuas
15	Semen, batu bata, Pasir, Cat, Thiner
16	cat, tiner, kayu
17	Besi, Semen, Paku, Kawat, Pasir
18	seng, paku, pvc, Pasir, batu bata
19	Kuas, Thiner
20	Batu Bata, Besi, Semen, Kawat, Pasir
21	Semen, Batu Bata, Kayu, Pasir
22	kawat, besi, batu bata, semen, pasir
23	Semen, Seng, Triplek, Paku
24	Kayu, seng, pvc
25	Besi, Pvc, Thiner, Batu Bata
26	triplek, kuas, cat
27	semen, Pasir
28	kawat, Besi, Kuas
29	Batu bata, semen, Pasir
30	Besi, Pvc, Semen, Kawat, Paku, Batu Bata
31	pasir, kayu, paku, semen
32	Pasir, Batu Bata
33	kuas, Cat, triplek
34	Besi
35	seng, cat, kuas
36	Besi, Triplek, Paku
37	pvc, cat, kuas
38	Seng, Paku, Semen, Cat, Thiner, Pasir

4.2.2. Inisialisasi Data

Untuk mempermudah melakukan perhitungan maka dilakukan langkah – langkah inialisasi untuk masing-masing item sebagai berikut :

- | | |
|----------------|-------------------|
| 1. Seng (A) | 8. Kuas (H) |
| 2. Semen (B) | 9. Tiner (I) |
| 3. Paku (C) | 10. PVC (J) |
| 4. Triplek (D) | 11. Pasir (K) |
| 5. Cat (E) | 12. Kayu (L) |
| 6. Besi (F) | 13. Batu Bata (M) |
| 7. Kawat (G) | |

Tabel 2. Data transaksi penjualan toko bangunan Gunung Saiyo yang telah diinisialisasi

Transaksi	Item yang dibeli
1	A, B, C, K, M
2	D, E
3	B, D, E, F, H, M
4	B, E, F, K
5	A, C
6	B, M
7	B, F, K
8	C, D, F, G
9	C, E, H, L
10	B, K, M
11	A, B, C, J, K
12	H, I
13	B, F, G, M
14	A, B, C, D, H, K
15	B, E, I, K, M
16	E, I, L
17	B, C, F, G, K
18	A, C, J, K, M
19	H, I
20	B, F, G, K, M
21	B, K, L, M
22	B, F, G, K, M
23	A, B, C, D
24	A, J, L
25	F, I, J, M
26	D, E, H
27	B, K
28	F, G, H
29	B, K, M
30	B, C, F, H, J, M
31	B, C, K, L
32	K, M
33	D, E, H
34	F
35	A, E, H
36	C, D, F
37	E, H, J
38	A, B, C, E, I, K

4.3. Pencarian Algoritma Apriori

Pada tahap ini dilakukan perhitungan manual dengan menggunakan *algoritma apriori* dengan cara menentukan *frequent itemset* pada setiap transaksi. Untuk melakukan perhitungan *frequent itemset* maka ditentukan terlebih dahulu nilai *minimum support* dan *minimum confidence*. Pada perhitungan kasus ini di tentukan nilai *minimum support* 0.25(25%) dan nilai *minimum confidence support* 0,75(75%). Langkah pertama yang harus dilakukan setelah menentukan nilai *minimum support* dan *minimum confidence* yaitu melakukan pembentukan *frequent itemset* untuk 1 *itemset*.

Tabel 3. Penjumlahan item pada setiap transaksi

Transaksi	A	B	C	D	E	F	G	H	I	J	K	L	M
1	1	1	1	0	0	0	0	0	0	0	1	0	1
2	0	0	0	1	1	0	0	0	0	0	0	0	0
3	0	1	0	1	1	1	0	1	0	0	1	0	1
4	0	1	0	0	1	1	0	0	0	0	1	0	0
5	1	0	1	0	0	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0	0	0	1
7	0	1	0	0	0	1	0	0	0	0	1	0	0
8	0	0	1	1	0	1	1	0	0	0	0	0	0
9	0	0	1	1	0	0	0	1	0	0	0	1	0
10	0	1	0	0	0	0	0	0	0	0	1	0	1
11	1	1	1	0	0	0	0	0	0	1	1	0	0
12	0	0	0	0	0	0	0	1	1	0	0	0	0
13	0	1	0	0	0	1	1	0	0	0	0	0	1
14	1	1	1	1	0	0	0	1	0	0	1	0	0
15	0	1	0	0	1	0	0	0	1	0	1	0	1
16	0	0	0	0	1	0	0	0	1	0	0	1	0
17	0	1	1	0	0	1	1	0	0	0	1	0	0
18	1	0	1	0	0	0	0	0	0	1	1	0	1
19	0	0	0	0	0	0	0	1	1	0	0	0	0
20	0	1	0	0	0	1	1	0	0	0	1	0	1
21	0	1	0	0	0	0	0	0	0	0	1	1	1
22	0	1	0	0	0	1	1	0	0	0	1	0	1
23	1	1	1	1	0	0	0	0	0	0	0	0	0
24	1	0	0	0	0	0	0	0	0	1	0	1	0
25	0	0	0	0	0	1	0	0	1	1	0	0	1
26	0	0	0	1	1	0	0	1	0	0	0	0	0
27	0	1	0	0	0	0	0	0	0	0	1	0	0
28	0	0	0	0	0	1	1	0	0	0	0	0	0
29	0	1	0	0	0	0	0	0	0	0	1	0	1
30	0	1	1	0	0	1	1	0	1	0	0	0	1
31	0	1	1	0	0	0	0	0	0	0	1	1	0
32	0	0	0	0	0	0	0	0	0	0	1	0	1
33	0	0	0	1	1	0	0	1	0	0	0	0	0
34	0	0	0	0	0	1	0	0	0	0	0	0	0
35	1	0	0	0	1	0	0	1	0	0	0	0	0
36	0	0	1	1	0	1	0	0	0	0	0	0	0
37	0	0	0	0	1	0	0	1	0	1	0	0	0
38	1	1	1	0	1	0	0	0	1	0	0	0	1
Jumlah	9	20	13	9	10	13	7	10	7	5	17	5	15

Setelah tabel 3 selesai maka dilakukanlah, langkah selanjutnya yaitu penentuan *frequent 1-itemset* seperti terlihat dalam tabel 4 di bawah ini :

Tabel 4. Perhitungan *frequent 1-itemset*

No	Item	Support
1	A	$(9/38) \times 100\% = 24\%$
2	B	$(20/38) \times 100\% = 53\%$
3	C	$(13/38) \times 100\% = 34\%$
4	D	$(9/38) \times 100\% = 24\%$
5	E	$(10/38) \times 100\% = 26\%$
6	F	$(13/38) \times 100\% = 34\%$
7	G	$(7/38) \times 100\% = 18\%$
8	H	$(10/38) \times 100\% = 26\%$
9	I	$(7/38) \times 100\% = 18\%$
10	J	$(5/38) \times 100\% = 13\%$
11	K	$(17/38) \times 100\% = 45\%$
12	L	$(5/38) \times 100\% = 13\%$
13	M	$(15/38) \times 100\% = 39\%$

Data di atas menggambarkan bentuk data satu *itemset*, yang terdiri atas *atribute item* sebagai nama *item* yang ada. Dari tabel 4.4 perhitungan *frequent 1-itemset* maka item yang memenuhi nilai *minimum support* 0.25(25%) dapat terlihat pada tabel 5 di bawah ini :

Tabel 5. frequent 1-itemset yang memenuhi nilai minimum support

No	Item	Support
1	B	$(20/38) \times 100\% = 53\%$
2	C	$(13/38) \times 100\% = 34\%$
3	E	$(10/38) \times 100\% = 26\%$
4	F	$(13/38) \times 100\% = 34\%$
5	H	$(10/38) \times 100\% = 26\%$
6	K	$(17/38) \times 100\% = 45\%$
7	M	$(15/38) \times 100\% = 39\%$

Tabel 5 di atas merupakan *frequent 1-itemset* yang memenuhi nilai *minimum support*. Di mana data diatas akan digunakan untuk membentuk pola kombinasi 2-*itemset* mulai dari pembentukan pola kombinasi 2-*itemset* sampai pada pembentukan pola kombinasi 3-*itemset*.

Pembentukan *frequent 2-itemset* dibentuk dengan mengkombinasikan item – item hasil perhitungan *frequent 1-itemset* yang memenuhi nilai *minimum support* yang telah kita tentukan. Kombinasi – kombinasi yang dibentuk dari *frequent 2-itemset* dapat dilihat pada tabel 6 di bawah ini :

Tabel 6. Perhitungan frequent 2-itemset

No	Item	Support
1	BC	$(8/38) \times 100\% = 21\%$
2	BE	$(3/38) \times 100\% = 8\%$
3	BF	$(8/38) \times 100\% = 21\%$
4	BH	$(2/38) \times 100\% = 5\%$
5	BK	$(15/38) \times 100\% = 39\%$
6	BM	$(12/38) \times 100\% = 32\%$
7	CE	$(1/38) \times 100\% = 3\%$
8	CF	$(4/38) \times 100\% = 11\%$
9	CH	$(2/38) \times 100\% = 5\%$
10	CK	$(6/38) \times 100\% = 16\%$
11	CM	$(4/38) \times 100\% = 11\%$
12	EF	$(2/38) \times 100\% = 5\%$
13	EH	$(5/38) \times 100\% = 13\%$
14	EK	$(3/38) \times 100\% = 8\%$
15	EM	$(3/38) \times 100\% = 8\%$
16	FH	$(2/38) \times 100\% = 5\%$
17	FK	$(6/38) \times 100\% = 16\%$
18	FM	$(6/38) \times 100\% = 16\%$
19	HK	$(2/38) \times 100\% = 5\%$
20	HM	$(1/38) \times 100\% = 3\%$
21	KM	$(10/38) \times 100\% = 26\%$

Tabel 6 perhitungan *frequent 2-itemset* di atas menggambarkan bentuk data kombinasi dua *itemset*. Dengan menetapkan nilai *minimum support* 0.25(25%)

, maka item yang memenuhi nilai *minimum support* dapat kita lihat pada tabel 7 di bawah ini :

Tabel 7. frequent 2-itemset yang memenuhi nilai minimum support

No	Item	Support
1	BK	$(15/38) \times 100\% = 39\%$
2	BM	$(12/38) \times 100\% = 32\%$
3	KM	$(10/38) \times 100\% = 26\%$

Data pada Tabel 7 adalah *frequent 2-itemset* yang memenuhi nilai *minimum support* yaitu BK, BM, dan KM. Setelah item *frequent 2-itemset* di dapatkan maka langkah selanjutnya yaitu melakukan pembentukan kombinasi *frequent 3-itemset*. Kombinasi – kombinasi yang terbentuk dapat kitalihat pada tabel 8 di bawah ini:

Tabel 8. Perhitungan frequent 3-itemset

No	Item	Support
1	BKM	$(8/38) \times 100\% = 21\%$

Tabel 8 menunjukkan bahwa hasil kombinasi *frequent 3-itemset* yang terbentuk dari kombinasi *frequent 2-itemset* yang memenuhi nilai *minimum support*. Dari gambaran tabel 8 maka item di atas tidak memenuhi nilai *minimum support* 0.25(25%) sehingga perhitungan *frequent itemset* dihentikan.

4.4. Pembentukan Rule Association

Setelah pencarian *algoritma apriori* selesai dilakukan, maka langkah selanjutnya yaitu mencari *rule association* yang memenuhi syarat *minimum confidence*. Pembentukan *rule association* dengan menghitung menggunakan aturan asosiasi A ke B. Kekuatan aturan asosiasi dilihat dengan cara melihat nilai *lift*nya.

Lift menunjukkan adanya tingkat kekuatan rule atas kejadian acak dari *antecedent* dan *consequence* berdasarkan pada *support*nya masing – masing. Hal ini akan mermbalikan informasi tentang perbaikan dan peningkatan probabilitas dari consequent berdasarkan antecedent. *Lift* didefinisikan sebagai berikut:

$$Lift = \frac{Confidence}{Confidence Benchmark}$$

Dimana:

$$Confidence Benchmark = \frac{\text{jumlah } h \text{ transaksi} \mid \text{items dalam consequent}}{\text{total jumlah } h \text{ transaksi}}$$

Jika nilai *lift* lebih besar dari pada 1 menunjukkan adanya manfaat dari aturan tersebut. Lebih tinggi nilai *lift* maka lebih besar kekuatan asosiasinya. Bila syarat *minimum confidence* adalah 0.75(75%) maka nilai *confidence*, nilai *lift* dari *frequent itemset* dapat kita lihat pada tabel 9 di bawah ini :

Tabel 9. Calon rule association

No	If Antecedent then Consequent	Lift	Confidence
1	If buy B then buy K	1,67	$(15/20) \times 100\% = 75\%$
2	If buy K then buy B	1.66	$(15/17) \times 100\% = 88\%$
3	If buy B then buy M	1.54	$(12/20) \times 100\% = 60\%$
4	If buy M then buy B	1.51	$(12/15) \times 100\% = 80\%$
5	If buy K then buy M	1.51	$(10/17) \times 100\% = 59\%$
6	If buy M then buy K	1.49	$(10/15) \times 100\% = 67\%$

Tabel 9 merupakan tabel calon *Rule Association* yang terdiri dari rule - rule yang dihasilkan dari pencarian *frequent itemset*. Dari rule - rule di atas yang memenuhi nilai *minimum confidence* dapat kita lihat pada tabel 10 di bawah ini:

Tabel 10. Rule association yang memenuhi minimum confidence

No	Rule	Lift	Confidance
1	If buy B then buy K	1.67	75 %
2	If buy K then buy B	1.66	88%
3	If buy M then buy B	1.51	80%

Tabel 10 menggambarkan *Rule Association* yang memenuhi nilai *minimum confidence*. Setelah selesai melakukan pencarian tersebut maka langkah selanjutnya yaitu menentukan item mana saja yang memenuhi nilai *minimum support* dan *minimum confidence* yang di gambarkan pada tabel 11 di bawah ini :

Tabel 11. Rule association nilai minimum support dan confidence

No	Rule	Lift	Support	Confidence
1	If buy B then buy K	1.67	39%	75 %
2	If buy K then buy B	1.66	39%	88%
3	If buy B then buy M	1.54	32%	60 %
4	If buy M then buy B	1.51	32%	80 %
5	If buy K then buy M	1.51	26%	59%
6	If buy M then buy K	1.49	26%	67%

Dari tabel diatas maka dapat kita lihat rule - rule yang memenuhi nilai *minimum support* dan *minimum confidence*. Hasil ini sangat penting karena pada tahap inilah yang menentukan sukses atau tidaknya langkah terakhir yaitu langkah penarikan kesimpulan.

4.5. Penarikan Kesimpulan

Setelah proses pencarian *Algoritma Apriori* dan pembentukan *Rule Association* telah selesai dilakukan, maka langkah terakhir yang dilakukan adalah penarikan kesimpulan terhadap rule - rule yang memenuhi nilai *minimum support* dan nilai *minimum confidence* berdasarkan tabel 11. Hasil dari rule - rule yang memenuhi nilai *minimum support* dan nilai

minimum confidence dapat terlihat pada tabel 12 dibawah ini :

Tabel 12. Rule Association yang memenuhi minimum support dan minimum confidence

No	Rule	Lift	Support	Confidance
1	If buy B then buy K	1.67	39 %	75 %
2	If buy K then buy B	1.66	39 %	88%
3	If buy M then buy B	1.51	32 %	80%

5. Implementasi dan Pengujian

5.1. Implementasi

Tahapan pengujian dengan *Tanagra 1.4* dilakukan dalam lingkungan berbasis windows yaitu dengan sistem operasi Windows 7, kemudian didukung dengan perangkat keras dan perangkat lunak yang menunjang. Tujuan dilakukannya pengujian adalah :

1. Untuk mengetahui apakah hasil analisa data secara manual akan sama dengan hasil analisa data dengan menggunakan perangkat lunak *Data Mining*, yaitu perangkat lunak *Tanagra 1.4*.
2. Untuk menguji bagaimana *association rule* akan dihasilkan jika menggunakan nilai *minimum support* dan *confident* yang berbeda-beda.

5.2. Pengujian

Pada tahap ini hal pertama yang dilakukan oleh penulis lebih berfokus kepada hal-hal sebagai berikut:

1. Pengujian terhadap hasil analisa data dan *rule - rule* yang terbentuk yang dilakukan melalui perhitungan secara manual apakah sama dengan hasil analisa data dengan menggunakan *Tanagra 1.4*.
2. Melakukan pengujian mengenai *association rule* yang terbentuk jika menggunakan *minimum support* dan *confident* yang berbeda-beda.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Transaksi	Seng	Semen	Palu	Triplek	Cat	Besi	Kawat	Kuas	Timor	PVC	Pasir	Kayu	Batu_Bata
2	1	1	1	1	0	0	0	0	0	0	0	0	1	0
3	2	0	0	0	1	1	0	0	0	0	0	0	0	0
4	3	0	1	0	1	1	1	0	1	0	0	1	0	1
5	4	0	1	0	0	1	1	0	0	0	0	1	0	0
6	5	1	0	1	0	0	0	0	0	0	0	0	0	0
7	6	0	1	0	0	0	0	0	0	0	0	0	0	1
8	7	0	1	0	0	0	1	0	0	0	0	1	0	0
9	8	0	0	1	1	0	1	1	0	0	0	0	0	0
10	9	0	0	1	1	0	0	1	0	0	0	0	1	0
11	10	0	1	0	0	0	0	0	0	0	0	1	0	1
12	11	1	1	1	0	0	0	0	0	0	1	1	0	0
13	12	0	0	0	0	0	0	0	1	1	0	0	0	0
14	13	0	1	0	0	0	1	1	0	0	0	0	0	1
15	14	1	1	1	1	0	0	0	1	0	0	1	0	0
16	15	0	1	0	0	1	0	0	0	1	0	1	0	1

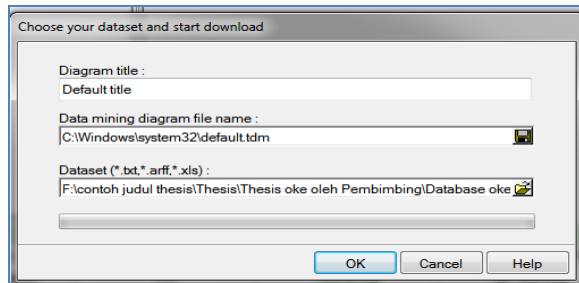
Gambar 2. Dataset dalam bentuk format (.xls)

Setelah tahap diatas dilakukan selanjutnya penulis menyiapkan *dataset* (data transaksi yang akan diolah), dataset tersebut dikonversikan ke dalam format file

microsoft spreadsheet (.xls) atau ke dalam format text (.txt). Dalam pengujian ini penulis mengkonversikan ke dalam format (.xls), bentuk format *dataset* yang digunakan dapat digambarkan sebagai gambar 2 di atas:

a. Import Dataset

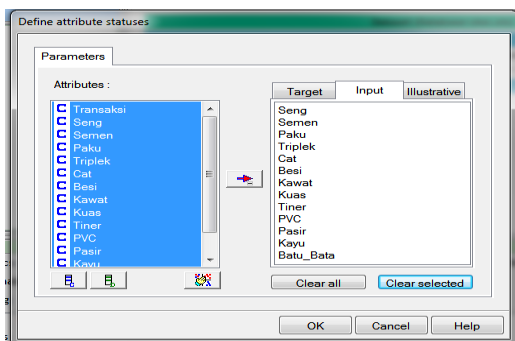
Dataset yang akan digunakan harus sudah dalam bentuk *1-itemset*, seperti pada tabel Kemudian *dataset* tersebut akan di *import* ke dalam *Tanagra 1.4* untuk diproses selanjutnya. Dengan cara seperti gambar 3 :



Gambar 3. Import dataset ke tanagra 1.4

b. Mendefinisikan Atribut Untuk Analisis

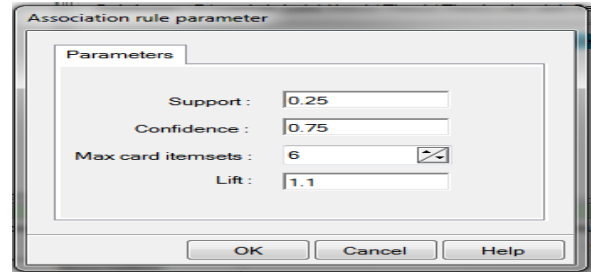
Setelah mengimpor *dataset*, berarti komponen *association* telah ditambahkan ke dalam diagram. Selanjutnya kita akan menentukan atribut-atribut yang akan digunakan sebagai input untuk mendapatkan *association rule* nantinya. Langkah menentukan atribut-atribut yang akan digunakan sebagai input adalah memilih *define attribute statuses* seperti gambar 4. berikut:



Gambar 4. Menginputkan atribut yang sudah dipilih

c. Algoritma Apriori

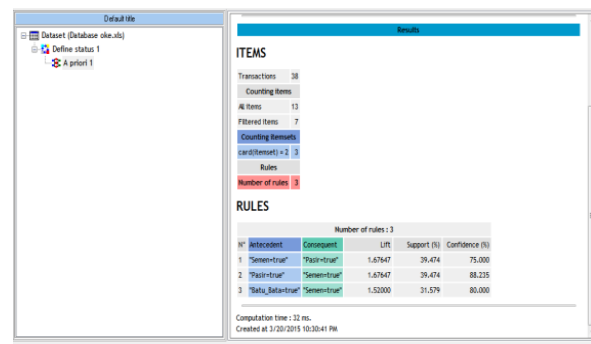
Langkah selanjutnya adalah menggunakan *algoritma apriori* untuk memproses (analisis) *dataset* yang telah di-*import* untuk mendapatkan *frequent itemset*, yaitu dengan memilih pilihan *algoritma apriori* pada tab “*Association*”. Digambarkan sebagai berikut:



Gambar 5. Menentukan support dan confidence

d. Mendapatkan Rule

Pada tahap perancangan penulis menggunakan *minimum support* 0.25 (25%) dan *minimum confidence* 0.75 (75%), sedangkan untuk maksimum *itemset* tidak ditentukan. Kemudian untuk memproses dan menampilkan *rule* yang memenuhi dapat dilakukan dengan cara klik kanan komponen “*A priori 1*” dan pilih menu “*Execute*”, lalu klik kanan komponen dan pilih menu “*View*”. Setelah itu perangkat lunak *Tanagra 1.4* akan memproses, dan kemudian akan menampilkan *rule - rule* yang memenuhi *minimum support* dan *minimum confidence*. Dapat digambarkan pada gambar 6 berikut:



Gambar 6. Association rule yang diperoleh

Berdasarkan *association rule* yang diperoleh melalui proses analisa dengan perangkat lunak *Tanagra 1.4* di atas, diketahui bahwa hasil tersebut sama dengan hasil perhitungan yang penulis lakukan pada tahap perancangan, yaitu terdapat 3(tiga) buah *rule*.

5.3. Pembahasan dan Hasil Pengujian

Dari dua hasil pengujian yang telah dilakukan yaitu proses secara manual dan menggunakan software *Tanagra 1.4* di atas, diketahui bahwa hasil tersebut sama dengan hasil perhitungan yang penulis lakukan pada tahap perancangan, yaitu terdapat 3(tiga) buah *rule*.

Rule - rule yang dihasilkan dari proses manual dan software *Tanagra 1.4* yaitu :

1. Jika membeli Semen maka akan membeli Pasir dengan *support* 39% dan *confidence* 75%.
2. Jika membeli Pasir maka akan membeli Semen dengan *support* 39% dan *confidence* 88%.
3. Jika membeli Batu_Bata maka akan membeli Semen dengan *support* 32% dan *confidence* 80%.

6. Kesimpulan

Kesimpulan yang dapat diambil dalam penelitian ini seperti yang telah diuraikan pada bab-bab yang sudah dibahas sebelumnya yaitu :

1. Metode *Market Basket Analysis (Rule Association)* dengan *Algoritma Apriori* ini dapat digunakan untuk membantu Toko Bangunan Gunung Saiyo mengetahui pola beli konsumen dan item yang sering dibeli oleh konsumen.
2. Metode *Market Basket Analysis (Rule Association)* dengan *Algoritma Apriori* ini dapat meningkatkan pelayanan dan penjualan barang pada Toko Bangunan Gunung Saiyo.
3. Hasil dari proses pencarian dengan software *Tanagra 1.4* membuktikan bahwa hubungan yang terjadi antar *item* sangat penting dan kuat, karena ada pembelian satu *item* terkait pada *item* lainnya.

Referensi

- [1] Johan. (2010). "Analisa Keranjang Pasar dengan Algoritma Apriori pad Data Transaksi Mini Market Lima Bintang"
- [2] Sheenu Verma. (2014). "An Efective Dynamic Unsupervised Clustering Algorithmic Appoach for Market Basket Analysis"
- [3] Fadlina. (2014). " Data Mining Untuk Analisa Tingkat Kejahatan Jalanan dengan Algoritma Association Rule Metode Apriori"
- [4] M.A. Hambali. (2014) "Data Mining in Market Basket Transaction : An Association Rule Mining Approach"
- [5] Musungwini, Samuel, Zhou, Tinashe Gwendolyn. (2014) "An Investigation of the Role of Product, Place, Promotion, and Price (4Ps) in Market Basket Analysis: A Case Study of Retail Shops in Gweru Zimbabwe"
- [6] Harpreet Kaur dan Kawaljeet Singh. (2013). " Market Basket Analysis of Sport Store using Association Rules"
- [7] Dr.M. Dhanabhakym dan Dr.M. Punithavalli. (2013). "An Efficient Market Basket Analysis Based on Adaptive Association Rule Mining With Faster Rule Generation Alghorithm"
- [8] Goldie Gunadi dan Dana Indra Sensuse. (2012). "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-Growth) Studi Kasus Percetakan PT Gramedia"
- [9] Arief Jananto. (2012). "Penggunaan Market Basket Analysis untuk Menentukan PolaKompetensi Mahasiswa"
- [10] Jongwook Woo. (2012). "Jurnal Penanganan Atribut Citra dengan Wavelet untuk Pengembangan Aloritma C4.5"
- [11] Mahendra Tiwari, Randhir Singh, Shivendra Kumar Singh. (2012). "Association Rule Mining Techniques A General Survey and Empirical Comparative Evaluation"
- [12] Dr.M. Dhanabhakym dan Dr.M. Punithavalli. (2011). "A Survey On Data Mining Algorithm for Market Basket Analysis"