

# Perancangan Data Mining dalam Analisis Asosiasi Kuantitatif Pembelian Item Barang dengan Metode Apriori

I Gusti Ayu Sri Melati, I Gusti Ayu Desi Saryanti

STMIK STIKOM Bali

Jl. Raya Puputan No. 86 Renon, (0361) 244445/(0361) 264773

e-mail: melati@stikom-bali.ac.id, desi@stikom-bali.ac.id

## Abstrak

Pada umumnya semua perusahaan mengumpulkan informasi melalui sistem database yang berguna untuk menampung data transaksi, kemudian nantinya data tersebut diolah sehingga dapat diketahui tingkat dan volume suatu penjualan, pembelian pada suatu waktu tertentu dan sebagainya, sehingga seorang pengambil keputusan dapat mengatur strategi pemasaran bagi perusahaannya. Dewasa ini muncul tuntutan bagi seorang pengambil keputusan, untuk dapat melihat peluang-peluang yang dapat meningkatkan penjualan di perusahaannya, berdasarkan informasi yang relevan. Pengambil keputusan perlu memakai suatu sistem yang dapat mendukungnya dalam mengambil keputusan secara cepat dan juga tepat. Informasi atas suatu produk dapat digunakan oleh perusahaan untuk meningkatkan penjualan. Salah satu cara untuk mendapatkan informasi adalah dengan melakukan data mining. Perancangan data mining ini akan menggunakan algoritma apriori untuk melakukan analisa keranjang pasar. Data yang diperlukan diambil dari data transaksi penjualan selama periode tertentu dan diolah sehingga menghasilkan frequent itemset. Sehingga pada akhirnya hasil yang didapatkan dari proses data mining ini dapat dijadikan suatu informasi baru yang dapat digunakan oleh sebuah perusahaan untuk meningkatkan produksi maupun penjualan produk-produknya

**Kata kunci:** Data mining, apriori

## Abstract

Generally all companies gather the information through a database system which is useful to contain the transaction data, and then later the data will be processed so that it can be known the level and volume of sales, the purchase of certain time etc, and therefore a decisionmaker can manage a marketing strategy for his company. Nowadays there is a demand for a decisionmaker to see through some chances that can improve sales in his company based on some relevant information. The decisionmaker needs to use a system which can support him in making the right decision precisely. Information of a product can be used by the company to increase sales. One way to get information is through data mining. The designing of this data mining will use apriori algorithm to do the analysing of market chart. The data needed will be taken from from the sales transaction data for a certin time of period and it will be processed to get the frequent itemset. At the end the result gotten from this data mining process can become a new information that can be usedby a company to increase production or sales of its products.

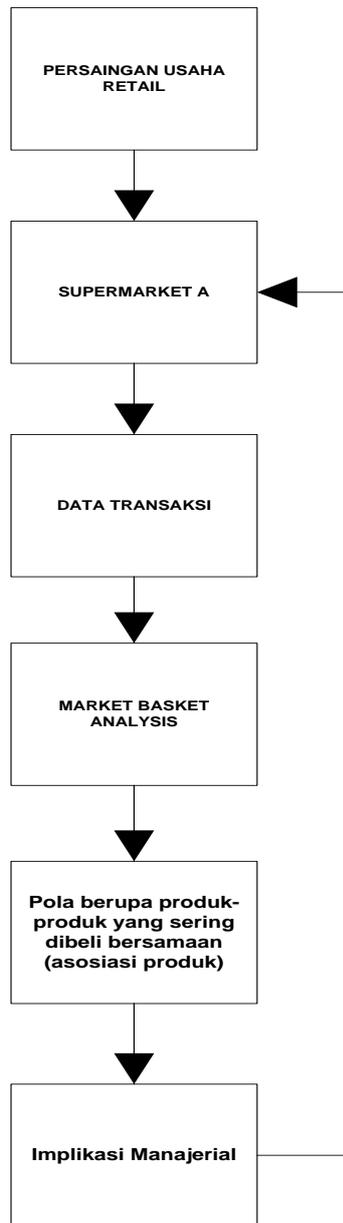
**Keywords:** Data mining, apriori

## 1. Pendahuluan

Pasar swalayan atau supermarket merupakan sarana atau tempat penyedia barang kebutuhan sehari-hari mulai dari sembilan bahan pokok sampai kebutuhan rumah tangga lainnya. Setiap harinya di supermarket tersebut terjadi beberapa transaksi penjualan. Data transaksi penjualan akan terus bertambah setiap harinya. Semakin banyak data transaksi yang tersimpan menyebabkan penyimpanan data yang sangat besar. Biasanya data-data transaksi penjualan tersebut hanya digunakan sebagai arsip dan tidak diketahui apa manfaat dari data tersebut untuk selanjutnya.

Di dalam kumpulan data yang sangat besar tersebut, memiliki informasi – informasi yang tersembunyi. Kumpulan data tersebut bisa diolah untuk menghasilkan informasi informasi yang bermanfaat yang bisa digunakan untuk pengambilan suatu keputusan dan untuk memperoleh pengetahuan. Pengolahan data tersebut bisa dilakukan dengan menggunakan teknik-teknik tertentu. Salah

satu teknik yang digunakan dalam pengolahan data tersebut adalah dengan menggunakan metode algoritma apriori.



Gambar 1. Kerangka Pemikiran Penelitian

Algoritma apriori adalah algoritma analisis keranjang pasar / market basket analysis yang digunakan untuk menghasilkan association rule, dengan pola “if then”.Market basket analysis merupakan salah satu teknik dari data mining yang mempelajari tentang perilaku kebiasaan konsumen dalam membeli barang secara bersamaan dalam satu waktu. Teknik tersebut bisa diterapkan dalam data yang sangat besar seperti data transaksi penjualan (marsela dan veronica, 2004).

Umumnya, pada analisis keranjang pasar diasumsikan data yang dimasukkan berupa atribut biner yang disebut item-item. Namun dalam kebanyakan domain bisnis dan data scientific mempunyai tipe atribut yang lebih beragam. Atribut bisa berbentuk kuantitatif ( seperti : produk, waktu pembelian ) atau kategori ( seperti: diskon ). Pertanyaan yang muncul, apakah ada hubungan antara produk yang dibeli dengan waktu pembelian produk serta diskon yang diberikan setelah para pelanggan berbelanja di supermarket. Jika ada hubungan positif, berarti dapat dikatakan analisis yang digunakan sudah baik, tetapi jika tidak maka dapat menjadi suatu masukan bagi evaluasi analisis yang ada. Dari hasil penelitian ini diharapkan dapat digali informasi-informasi yang mungkin berguna untuk melihat kecenderungan-

kecenderungan asosiasi yang ada antara produk , waktu pembelian dan diskon. Sehingga pada akhirnya hasil yang didapatkan dari proses data mining ini dapat dijadikan suatu informasi baru yang dapat digunakan oleh pemilik supermarket untuk meningkatkan produksi maupun penjualan produk-produknya.

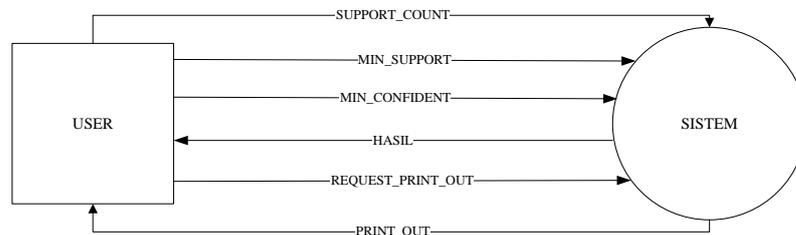
**2. Metode Penelitian**

Sejak berkembangnya pasar retail modern di Indonesia, semakin banyak ritel yang melakukan berbagai usaha demi mempertahankan usahanya. Salah satu upaya yang dilakukan untuk dapat mempertahankan bisnis ritel ini adalah dengan meningkatkan dan mempertahankan tingkat pertumbuhan penjualannya. Pendekatan yang dapat dilakukan dalam rangka meningkatkan volume penjualan tersebut adalah dengan memahami dan mengetahui pola perilaku belanja konsumen. Hampir semua supermarket menggunakan komputersisasi dalam penyimpanan data penjualan dan pembelian, sehingga akan dihasilkan banyak data transaksi.

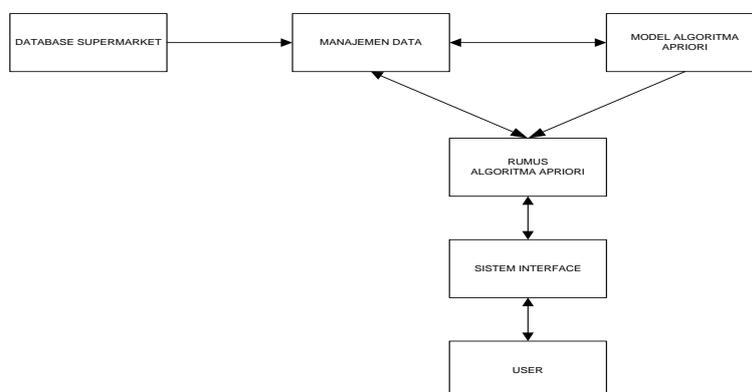
Data transaksi tersebut dapat digunakan sebagai sumber informasi penting dalam mempertahankan keberadaan usaha tersebut. Pengetahuan tentang pola perilaku belanja tersebut dapat diketahui dengan menggunakan analisis keranjang belanja (Market Basket Analysis). Analisis ini merupakan salah satu metode penambangan data (data mining) yang bertujuan untuk menemukan produk-produk yang sering dibeli bersamaan dari data transaksi, yang dapat dimanfaatkan untuk menyusun tata letak sedemikian rupa sehingga memberikan kemudahan bagi konsumen dalam melakukan pembelian, dan dapat meningkatkan kemungkinan munculnya keinginan membeli (impulse buying).

**3. Hasil dan Analisa**

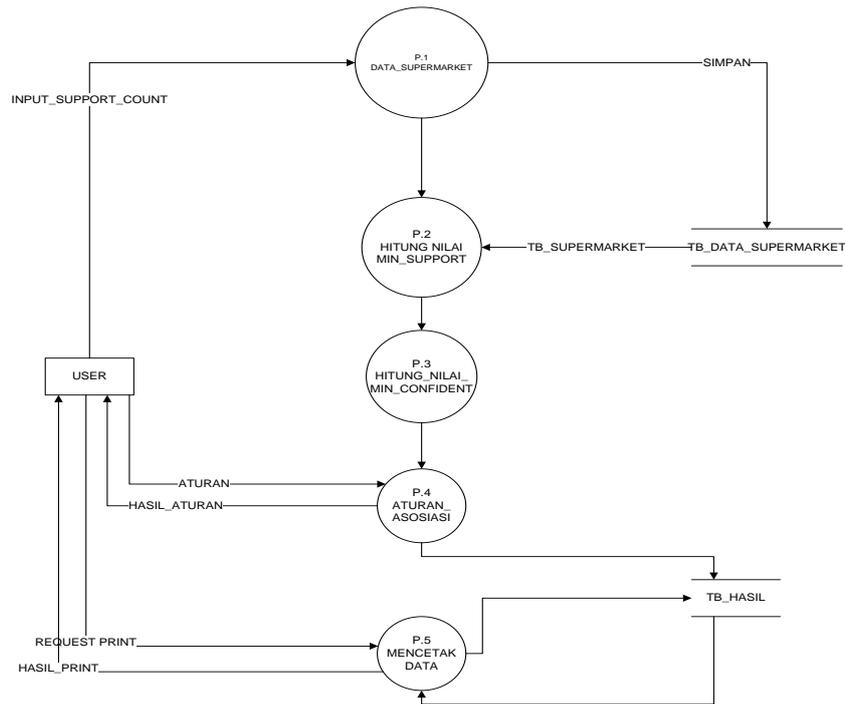
Perancangan diagram arus ini digunakan untuk menggambarkan alur data pada saat proses pemindahan data. Dalam proses pemindahan data ini yang dilakukan hanya perintah 1 arah, jadi proses akan berjalan secara otomatis setelah diberikan perintah oleh user melalui user interface yang sudah ada.



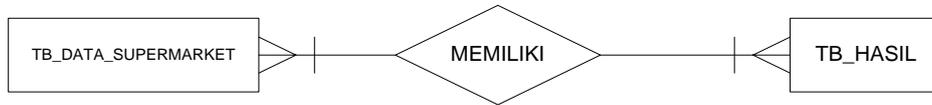
Gambar 2. Diagram Konteks perancangan Data Mining



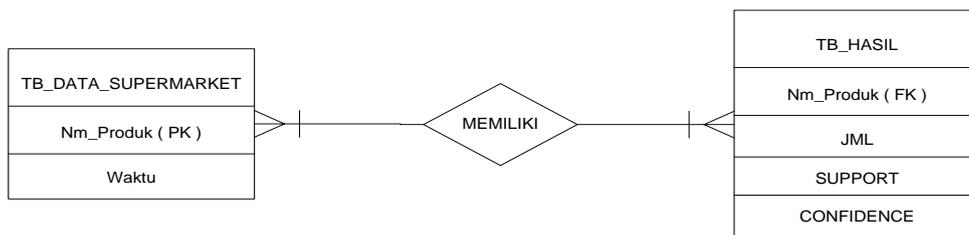
Gambar 3. Hirarki Chart perancangan Data Mining



Gambar 4. DFD Level 0 perancangan Data Mining

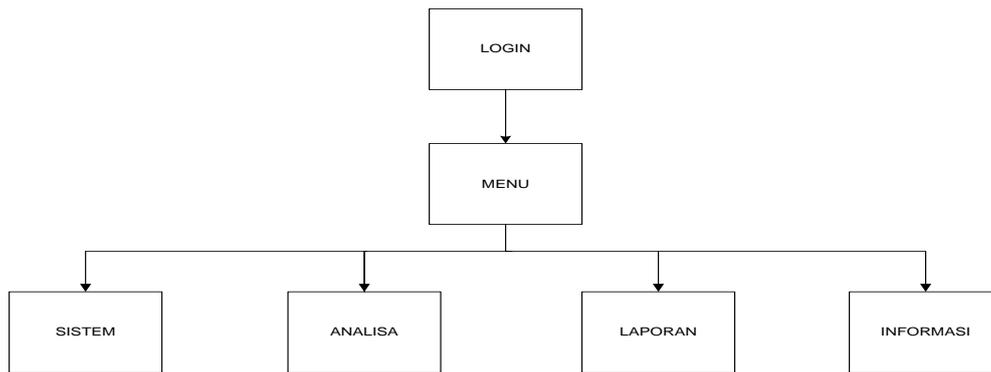


Gambar 5. ERD perancangan Data mining



Gambar 6. Konseptual Database

Pada diagram diatas terlihat primary key masing-masing tabel yang berelasi, suatu tabel dapat dikatakan berelasi apabila tabel yang direlasikan memiliki keterkaitan dalam hal ini pada primary key yang ada pada tabel.



Gambar 7. Struktur Menu

**3.1. Perancangan Antarmuka**

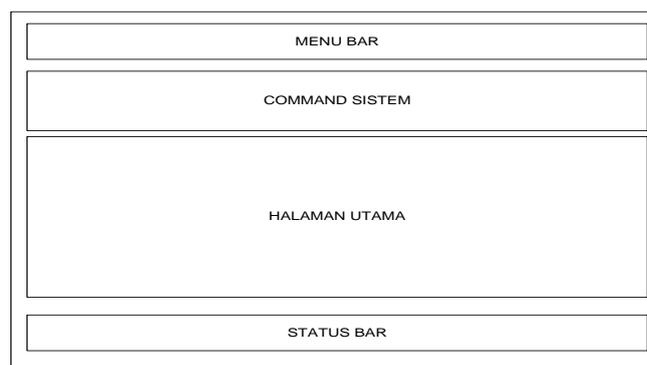
Perancangan Login untuk masuk ke menu utama, berikut desain loginnya :

USER NAME

PASSWORD

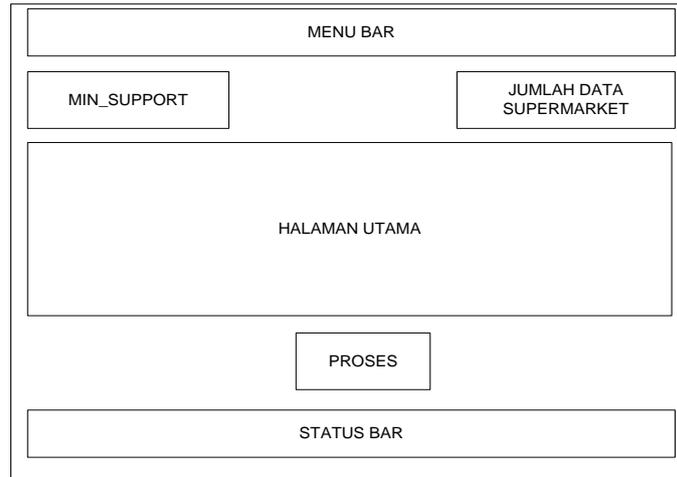
Gambar 8. Desain Login Awal

Desain utama yang merupakan rancangan awal untuk maintenance dan pengelolaan sistem. Berikut rancangan desainnya :



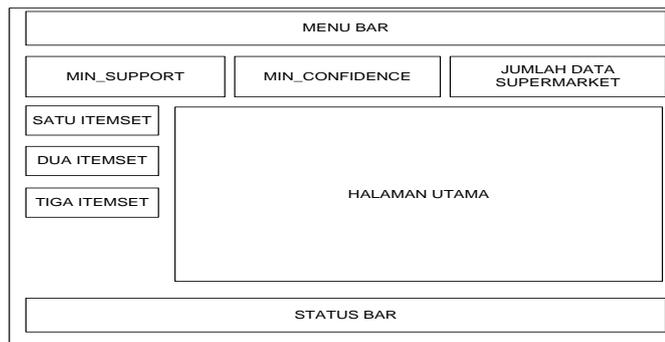
Gambar 9. Halaman Utama

Desain menu yang merupakan rancangan untuk memilih data Supermarket yang akan dianalisa. Berikut rancangan desainnya :



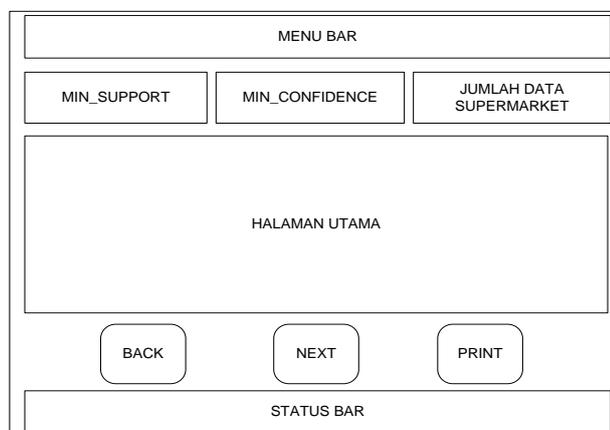
Gambar 10. Desain Menu Proses

Desain analisa merupakan rancangan untuk menganalisa data sehingga mendapatkan pola dari data tersebut. Berikut rancangan desainnya :



Gambar 11. Desain Analisa

Desain laporan analisa merupakan rancangan untuk melihat hasil data yang sudah dianalisa . Berikut rancangan desainnya :



Gambar 12. Desain Laporan Analisa

### 3.2. Hasil dan Pembahasan

#### *Kaidah Asosiasi pada Item-item Penjualan*

Analisis pertama yang dilakukan menggali pengaruh waktu terhadap penjualan suatu produk. Hasil analisis pengaruh waktu terhadap jumlah penjualan buah dan mie di Supermarket A dapat dilihat pada Tabel 1.

Tabel 1. Support\_count untuk kandidat 1 itemset

1 Itemset	Support Count	Support ( % )
buah	2089	49.63
mie	2120	50.37
t1	1481	35.19
t2	1145	27.20
t3	1583	37.61

Keterangan :

T1 : Waktu pembelian antara jam 09:00-13:06 WITA

T2 : Waktu pembelian antara jam 13:07-17:13 WITA

T3 : Waktu pembelian antara jam 17:14-21:20 WITA

Berdasarkan Tabel 1 terdapat nilai *support* yang menunjukkan persentasi *item* buah, mie, dan waktu pembelian yang muncul dari semua transaksi yang mengandung buah dan mie. Jumlah transaksi penjualannya sebanyak 4209 transaksi. Nilai *support* untuk buah adalah 49,63% yang berarti bahwa 49,63% dari seluruh transaksi yang terjadi mengandung *item* buah, sedangkan nilai *support* untuk mie adalah 50,37% yang berarti bahwa 50,37% dari seluruh transaksi yang terjadi mengandung *item* mie. Perbandingan nilai kedua *item* tersebut menunjukkan bahwa pembelian untuk *item* mie lebih banyak dari pada pembelian *item* buah. Konsumen berbelanja pada waktu yang berbeda-beda, dari tabel dapat terlihat Supermarket A banyak didatangi konsumen untuk berbelanja pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA), karena nilai *support* yang dihasilkan sebesar 37,61%. Jika dibandingkan dengan T2 (Waktu pembelian antara jam 13:07-17:13 WITA) 27,20%, dan T1 (Waktu pembelian antara jam 09:00-13:06 WITA) 35,19% karena nilai *support* keduanya lebih rendah.

Tabel dibawah, menunjukkan kandidat untuk *2itemset* dengan menentukan *support\_count*, untuk kandidat *frequent 2 itemset* (Tabel 2).

Tabel 2. Support count untuk kandidat 2 itemset produk dan waktu pembelian

2 Itemset	T1		T2		T3	
	Support count	Support (%)	Support count	Support (%)	Support count	Support (%)
Buah	730	17.34	651	14.61	744	17.68
Mie	751	17.84	530	12.60	839	19.93

Berdasarkan Tabel 2 terdapat nilai *support* waktu pembelian dengan buah dan nilai *support* waktu pembelian dengan mie. Nilai *support* yang dihasilkan antara waktu pembelian dengan buah yang paling tinggi adalah buah banyak dibeli pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) dengan nilai *support* 17,68% yang berarti bahwa 17,68% dari seluruh transaksi yang terjadi mengandung pembelian buah yang dilakukan pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA). Pembelian untuk mie banyak juga dilakukan pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) dengan nilai *support*

19,93%, yang berarti bahwa 19,93% dari seluruh transaksi yang terjadi mengandung pembelian mie pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA).

Menentukan kandidat untuk 3 *itemset* dengan menentukan *support\_count*, yang akan memperlihatkan ketiga *item* yang saling berasosiasi (Tabel 3).

Tabel 3. Support count untuk kandidat 3 itemset T1

3 Itemset ( t1 )	Buah		Mie	
	Support count	Support ( % )	Support count	Support ( % )
Buah	593	40.04	137	9.25
Mie	137	9.25	614	41.46

Berdasarkan Tabel 3 terdapat nilai *support* untuk waktu pembelian antara buah dan mie dibeli bersamaan (tidak memperhatikan urutan pembelian *item*) memiliki nilai *support* terkecil yaitu 9,25% (nilai *support* 9,25% dan nilai *support* 9,25%) yang berarti 9,25% dari semua transaksi yang terjadi mengandung buah dan mie yang dilakukan pada T1 (Waktu pembelian antara jam 09:00-13:06 WITA), sedangkan untuk nilai *support* yang paling tinggi adalah pada pembelian mie pada T1 (Waktu pembelian antara jam 09:00-13:06 WITA), sebesar 41,46% dibandingkan dengan pembelian buah pada T1 (Waktu pembelian antara jam 09:00-13:06 WITA), sebesar 40,04%.

Tabel 4. Support count untuk kandidat 3 itemset T2

3 Itemset ( t2 )	Buah		Mie	
	Support count	Support ( % )	Support count	Support ( % )
Buah	509	44.45	106	9.26
Mie	106	9.26	424	37.03

Berdasarkan Tabel 4 terdapat nilai *support* terkecil untuk waktu pembelian antara buah dan mie dibeli bersamaan (tidak memperhatikan urutan pembelian *item*) yaitu 9,26% (nilai *support* 9,26% dan nilai *support* 9,26%) yang berarti 9,26% dari semua transaksi yang terjadi mengandung buah dan mie yang dilakukan pada T2 (Waktu pembelian antara jam 13:07-17:13 WITA). Sedangkan untuk nilai *support* yang paling tinggi adalah pada pembelian buah pada T2 (Waktu pembelian antara jam 13:07-17:13 WITA), sebesar 44,45% dibandingkan dengan pembelian mie pada T2 (Waktu pembelian antara jam 13:07-17:13 WITA), sebesar 37,03%.

Tabel 5. Support count untuk kandidat 3 itemset T3

3 itemset ( T3 )	Buah		Mie	
	Support Count	Support ( % )	Support Count	Support ( % )
Buah	588	37.140	156	9.855
Mie	156	9.855	683	43.150

Berdasarkan Tabel 5 terdapat nilai *support* terkecil untuk waktu pembelian antara buah dan mie dibeli bersamaan (tidak memperhatikan urutan pembelian *item*) yaitu 9,855% ( nilai *support* 9,855% dan nilai *support* 9,855%) yang berarti 9,71% dari semua transaksi yang terjadi mengandung buah dan mie yang dilakukan pada T3 (Waktu pembelian antara jam 13:07-17:13 WITA). Sedangkan untuk nilai *support* yang paling tinggi adalah pada pembelian mie pada T3 (Waktu pembelian antara jam 13:07-17:13 WITA), sebesar 43,150% dibandingkan dengan pembelian buah pada T3 (Waktu pembelian antara jam 13:07-17:13 WITA), sebesar 37,140%.

Tabel 6. Nilai support dan confidence dari kaidah asosiasi

<i>Frequent k-itemset</i>	<i>Kaidah Asosiasi</i>	<i>Support (%)</i>	<i>Confidence (%)</i>
T1_Mie	T1→Mie	17,84	50,71
T2_Mie	T2→Mie	12,60	46,29
T3_Mie	T3→Mie	19,93	53,00
T1_Buah	T1→Buah	17,34	49,29
T2_Buah	T2→Buah	14,61	53,71
T3_Buah	T3→Buah	17,68	47,00
Mie_buah_T1	Mie dan buah → pagi	9,250	34,335
Mie_buah_T2	Mie dan buah → siang	9,260	26,565
Mie_buah_T3	Mie dan buah → malam	9,855	39,100

Berdasarkan Tabel 5.6 menunjukkan buah lebih banyak dibeli pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) 17,68% dibandingkan dengan T1 (Waktu pembelian antara jam 09:00-13:06 WITA) 14,61% dan T2 (Waktu pembelian antara jam 13:07-17:13 WITA) 17,34% dengan tingkat kepercayaan terbesar 49,29%. Kaidah tersebut dapat diterjemahkan bahwa, sebanyak 17,68% dari seluruh data, jika pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) maka buah dibeli, dengan derajat kepercayaan sebesar 49,29%. Mie lebih banyak dibeli pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) 19,93% dibandingkan dengan T2 (Waktu pembelian antara jam 13:07-17:13 WITA) 12,60% dan T1 (Waktu pembelian antara jam 09:00-13:06 WITA) 17,84% dengan tingkat kepercayaan terbesar 53,00%. Kaidah tersebut dapat diterjemahkan, bahwa sebanyak 19,93% dari seluruh data, jika pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) maka buah dibeli, dengan derajat kepercayaan sebesar 53,00%. Penelitian ini memberikan informasi bahwa pembelian buah dan mie lebih banyak dilakukan pada T3 (Waktu pembelian antara jam 17:14-21:20 WITA) sehingga pihak Supermarket lebih memperhatikan untuk persediaan produk yang dipengaruhi oleh waktu.

#### 4. Kesimpulan

Penelitian ini memperlihatkan bahwa Analisis Keranjang Belanja tidak hanya digunakan untuk menentukan kecenderungan pelanggan dalam membeli barang-barang secara bersamaan, tetapi dapat digunakan untuk melihat kapan suatu barang dibeli pada jumlah yang paling besar, mencari hubungan diantara himpunan barang-barang yang berbeda, dan memberikan informasi lain yang dihasilkan dalam transaksi penjualan. Hasil analisis ini memiliki informasi penting yang dapat meningkatkan kualitas kerja karyawan karena karyawan mengetahui kapan saja persediaan harus diperbanyak sehingga karyawan mempunyai perkiraan yang tepat dalam mempersiapkan produk-produk tersebut serta memberikan ide-ide bagi karyawan bagian promosi dengan melihat produk yang sering dibeli bersamaan.

#### References

- [1] Pramudiono, I. 2007. *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*. <http://www.ilmukomputer.org/wp-content/uploads/2006/08/iko-datamining.zip>
- [2] Kusriani, dan Emha Taufik Luthfi, 2009, “*Algoritma Data Mining*”, Penerbit Andi, Yogyakarta. Pramudiono, I., 2007, *Algoritma Apriori*, <http://datamining.japati.net/cgi-bin/indodm.cgi?bacaarsip&1172210143>
- [3] Santosa, Budi, 2007, “*Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*”, Graha Ilmu, Yogyakarta.
- [4] Prasetyo, Bowo, 2006, *Analisis Perilaku Pengunjung Menggunakan Data mining*, [www.beritaiptek.com](http://www.beritaiptek.com).
- [5] Han, Jiawei dan Kamber, Micheline, 2001, *Data mining Concepts and Techniques*, Academic Press, San Diego
- [6] Tan, Pang-Ning, Steinbach, Michael dan Kumar, Vipin, 2006 *Introduction to Data mining*, Pearson Education, Inc.