

PERBANDINGAN ASSOSSIATION RULE BERBENTUK BINER DAN FUZZY C-PARTITION PADA ANALISIS MARKET BASKET DALAM DATA MINING

Altien J. Rindengan¹⁾

¹⁾ Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Sam Ratulangi, Manado
Jl. Kampus Unsrat Manado, 95115
e-mail : altien@unsrat.ac.id

ABSTRAK

Salah satu analisis dalam *data mining* adalah *market basket analysis* untuk menganalisa kecenderungan pembelian suatu barang yang berasosiasi dengan barang yang lain. Dalam tulisan ini membahas aturan asosiasinya dengan mempertimbangkan jumlah *item* barang yang dibeli dalam satu transaksi. Asumsinya adalah keterkaitan pembelian suatu barang dengan barang yang lain dalam satu transaksi akan semakin kecil jika jumlah item barang yang dibeli semakin banyak. Tulisan ini menganalisa asosisasi antar *item* barang dengan membuat tabel transaksi dalam bentuk nilai *fuzzy set* dibandingkan dengan analisa asosiasi yang biasa dilakukan dalam bentuk biner. Berdasarkan analisis terhadap data yang digunakan memberikan hasil *support* dan *confidence* yang cenderung lebih kecil tetapi lebih realistis dibanding aturan asosisasi biasa.

Keywords: analisis *market basket*, *association rule*, *data mining*, *fuzzy c-partition*.

COMPARISON OF ASSOCIATION RULE WITH BINARY AND FUZZY C-PARTITION FORM AT MARKET BASKET ANALYSIS ON DATA MINING

ABSTRACT

One analysis in data mining is market basket analysis to analyze the purchase of a good trends associated with other items. In this paper discussing the association rules by considering the number of items purchased in one transaction. The assumption is that the purchase of a good relationship with the other items in one transaction will be smaller if the number of items purchased items more and more. This paper analyzes the association between the items of goods by making the transaction table in the form of fuzzy sets of values to compare with analysis of the usual associations in binary form. Based on the analysis of the data used to support and confidence of which tend to be smaller but more realistic than usual asosisasi rules.

Keywords: market basket analysis, association rule, data mining, fuzzy c-partition.

PENDAHULUAN

Market basket adalah istilah untuk menggambarkan kelompok barang yang cenderung dibeli oleh seorang konsumen dalam satu transaksi (misalkan di suatu supermarket). Misalkan kecenderungan membeli barang *X* jika ia membeli barang *Y*, demikian sebaliknya. Dari sisi pengelola supermarket bisa diperoleh informasi untuk mengantisipasi hal ini, misalkan dengan

menempatkan kedua barang tersebut pada tempat yang saling berdekatan agar mudah diperoleh konsumen.

Dalam *data mining*, metode yang biasa digunakan untuk menganalisis market basket adalah metode apriori. Metode ini adalah suatu algoritma dalam pencarian frekuensi barang dengan *association rule*, dimana metode ini cenderung menfokuskan pada hubungan antar barang dalam keseluruhan

transaksi saja tanpa memperhatikan hubungan tiap barang dalam tiap transaksi. Misalkan hubungan antara barang *X* dan *Y* hanya dihitung berapa kali keduanya muncul dalam keseluruhan transaksi tanpa memperhatikan jumlah barang yang terdapat dalam transaksi yang terdapat kedua barang tersebut. Akan tetapi, ada kemungkinan pembelian barang *X* atau *Y* tersebut dipengaruhi oleh pembelian barang yang lain sehingga akan memberikan hasil yang kurang akurat. Oleh karena itu perlu dianalisa hubungan antar barang yang bisa dilakukan dengan membuat tabel transaksi dalam bentuk *fuzzy c-partition*. Dimana metode ini dipakai untuk mengklasifikasi elemen-elemen dari suatu himpunan secara umum menjadi partisi-partisi berupa himpunan *fuzzy*. Konsep yang dilihat adalah bahwa suatu elemen akan memiliki hubungan/*similarity* dengan elemen lain jika ada dalam kelompok/class yang sama. Elemen-elemen tersebut akan memiliki hubungan yang lebih kuat jika sering berada dalam kelompok yang sama, tetapi dengan semakin banyaknya jumlah elemen pada suatu kelompok akan mengurangi derajat *similarity* antar tiap elemen dalam kelompok tersebut.

Beberapa penelitian sebelumnya yang berhubungan dengan tulisan ini adalah tentang alat analisis untuk *market basket analysis* (Auslender, 2004), *fuzzy association rule* (Intan, 2006 dan Helm, 2007), dan derajat kesamaan dalam *fuzzy partition* (Intan, 2002).

Tujuan dari tulisan ini adalah untuk mengetahui tingkat asosiasi antar barang dalam tabel transaksi yang dinyatakan dalam bentuk *fuzzy c-partition*. Ini akan berguna untuk membantu pengambil keputusan dalam menentukan kebijakan pemasaran di lingkungan usahanya.

TINJAUAN PUSTAKA

Market Basket Analysis

Market basket analysis adalah salah satu cara yang digunakan untuk menganalisis

data penjualan dari suatu perusahaan. Proses ini menganalisis perilaku pembelian konsumen dengan melihat asosiasi antar item-item yang berbeda pada tiap transaksi pembelian (Han dan Kamber, 2006). Dengan melihat hubungan kegunaan item-item, beberapa diantaranya mudah diperkirakan *item* apa saja yang dibeli secara bersamaan, misalnya kopi dan gula, sehinggaantisipasi ketersediaan kopi akan sama dengan gula. Namun mungkin saja ada pola pembelian *item* yang tidak diperkirakan sebelumnya, misalnya gula dan sabun yang tidak mempunyai hubungan sebagai barang substitusi atau komplementer, sehingga akan terjadi kesalahan dalam mengantisipasi ketersediaan gula atau sabun tersebut. Inilah salah satu manfaat yang dapat diperoleh dari melakukan *market basket analysis*. Dengan melakukan analisa ini, seorang manajer tidak perlu mengalami kesulitan untuk menemukan pola pembelian *item* apa saja yang mungkin dibeli secara bersamaan, karena data dari transaksi penjualan akan memberitahukannya sendiri. Dalam *data mining*, analisa ini dilakukan dengan *association rule*.

Association Rule

Association rule mining adalah suatu prosedur untuk mencari hubungan antar *item* dalam suatu data set yang ditentukan (Han dan Kamber, 2006). *Association rule* meliputi dua tahap yaitu mencari kombinasi yang paling sering terjadi dari suatu *itemset* dan mendefinisikan *condition* dan *result* (untuk *conditional association rule*). Masalah mendasar dari analisis asosiasi adalah bagaimana menemukan kaidah dalam bentuk $set_1 \rightarrow set_2$ (Possas, Wagner, Marcio dan Rodolfo, 2000). Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu, yaitu:

- *support*: suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu

item/itemset dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu *item/itemset* layak untuk dicari *confidence*-nya (misal, dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi yang menunjukkan bahwa *item A* dan *B* dibeli bersamaan).

- *confidence*: suatu ukuran yang menunjukkan hubungan antar 2 *item* secara *conditional* (misal, seberapa sering *item B* dibeli jika orang membeli *item A*).

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

Definisi formal untuk kedua ukuran ini dapat dirumuskan:

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

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

(Tan, Steinbach dan Kumar, 2006)

Kedua ukuran ini nantinya berguna dalam menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh *user*. Batasan tersebut umumnya terdiri dari *min_support* dan *min_confidence*. Contoh (suatu *association rule*), *If A then B* [*support=3%, confidence=70%*], dimana *A* dan *B* adalah kumpulan *item* yang dibeli oleh konsumen perusahaan Z. Artinya: *item A* dan *B* dibeli bersamaan sebesar 3% dari keseluruhan data transaksi yang dianalisis dan 70% dari semua konsumen yang membeli *item A* juga membeli *item B*. Dari

contoh di atas, jika *support*-nya $\geq \text{min_support}$ dan *confidence*-nya $\geq \text{min_confidence}$, maka *rule* tersebut bisa dikatakan sebagai *interesting rule*.

Algoritma Apriori

Dalam Han dan Kamber, 2006, didefinisikan bahwa “*apriori is an influential algorithm for mining frequent itemsets for Boolean association rules*”. Apriori adalah suatu algoritma yang sudah sangat dikenal dalam melakukan pencarian *frequent itemset* dengan *association rule*. Sesuai dengan namanya, algoritma ini menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Algoritma inilah yang biasanya dipakai dalam proses *data mining* untuk *market basket analysis*. Algoritma apriori memakai pendekatan *iterative (level-wise search)*, dimana *k-itemset* dipakai untuk menyelidiki *(k+1)-itemset*. Langkah-langkah dari algoritma ini adalah sebagai berikut (Ulmer, 2002):

1. *Set k=1* (menunjuk pada *itemset* ke-1).
2. Hitung semua *k-itemset* (*itemset* yang mempunyai *k-item*).
3. Hitung *support* dari semua calon *itemset*. Pilih *itemset* tersebut berdasarkan perhitungan *minimum support*.
4. Gabungkan semua *k-sized itemset* untuk menghasilkan calon *itemset k+1*.
5. *Set k=k+1*.
6. Ulangi langkah 3-5 sampai tidak ada *itemset* yang lebih besar yang dapat dibentuk.
7. Buat *final set* dari *itemset* dengan menciptakan suatu *union* dari semua *k-itemset*.

Contohnya, jika ada 10 transaksi (T1,T2,...,T10) untuk 5 *item* barang (B1,B2,...,B5) seperti pada Tabel 1, maka :

Tabel 1 Contoh 1 transaksi dengan 5 item barang

	B1	B2	B3	B4	B5
T1	1	1	0	1	1
T2	1	1	1	0	1
T3	0	0	1	0	0
T4	1	1	0	1	1
T5	1	1	0	1	1
T6	0	1	0	0	1
T7	1	1	0	0	1
T8	0	0	0	1	1
T9	1	1	1	0	1
T10	0	1	1	1	0

Support untuk 1-item set adalah:

$support(B1)=60\%$,

$support(B2)=80\%$,

$support(B3)=40\%$,

$support(B4)=50\%$,

$support(B5)=80\%$.

Jika ditetapkan $min_support=60\%$, maka yang memenuhi hanya B1, B2, dan B3.

Selanjutnya untuk 2-item set hanya dianalisa pada kombinasi 1-item set yang memenuhi $min_support$ yaitu kombinasi dari B1, B2, dan B3. Jika dihitung, diperoleh:

$support(B1 \cup B2)=60\%$,

$support(B1 \cup B5)=60\%$,

$support(B2 \cup B5)=70\%$.

Jika $min_support$ sama, maka ketiga kombinasi ini memenuhi. Untuk 3-item set, diperoleh $support(B1 \cup B2 \cup B5)=60\%$ yang memenuhi $min_support$. Jika selanjutnya dihitung nilai *confidence*-nya dengan menetapkan $min_confidence$ sebesar 80% maka diperoleh *interesting rule* sebagai berikut:

If B1 than B2 ($support=60\%$, $confidence=100\%$)

If B1 than B5 ($support=60\%$, $confidence=100\%$)

If B2 than B5 ($support=70\%$, $confidence=87.5\%$)

If B5 than B2 ($support=70\%$, $confidence=87.5\%$)

If B1 than (B2 \cup B5)

($support=60\%$, $confidence=100\%$)

If (B1 \cup B2) than B5

($support=60\%$, $confidence=100\%$)

If (B1 \cup B5) than B2

($support=60\%$, $confidence=100\%$)

If (B2 \cup B5) than B1

($support=60\%$, $confidence=87.5\%$).

Fuzzy c-Partition

Menurut (Klir dan Yuan, 2001), *fuzzy c-partition* dapat didefinisikan sebagai berikut:

Misalkan $I = \{i_1, i_2, \dots, i_n\}$ adalah domain dari data. *Fuzzy c-partition* dari I adalah *fuzzy subset* atau *fuzzy classes* dari $T = \{t_1, t_2, \dots, t_n\}$, yang memenuhi:

$$\sum_{m=1}^c \mu_{t_m}(i_k) = 1 \quad (1)$$

untuk semua $k \in N_n$

$$0 < \sum_{k=1}^c \mu_{t_m}(i_k) < n \quad (2)$$

untuk semua $m \in N_c$

dimana $\mu_{t_m}(i_k) \in [0,1]$

Fuzzy c-Covering

Fuzzy c-covering merupakan salah satu metode yang dipakai untuk mengklasifikasikan elemen-elemen dari suatu himpunan universal menjadi partisi-partisi berupa *fuzzy sets*. *Fuzzy c-covering* merupakan generalisasi dari metode *fuzzy c-partition* pada persamaan (1) menjadi :

$$\sum_{m=1}^c \mu_{t_m}(i_k) \geq 1 \quad (3)$$

untuk semua $k \in N_n$.

Degree of Similarity in Fuzzy c-Covering

Setiap *crisp element* dari data pada I juga dapat direpresentasikan sebagai *fuzzy set* dari *fuzzy classes* (T) dengan rumus (Intan dan Masao, 2003):

$$\mu_{i_m}(t_j) = \frac{\mu_{t_j}(i_m)}{\sum_{i \in I} \mu_{t_j}(i)} \quad (4)$$

dimana $\mu_{t_j}(i_m)$ adalah nilai *fuzzy partisi* t_j pada elemen i_m dan $\sum_{i \in I} \mu_{t_j}(i)$ adalah penjumlahan nilai *fuzzy partisi* t_j pada semua elemen dalam domain I .

Dalam model pembelian barang, t_j berhubungan dengan nomor transaksi (transaksi ke- j) dan elemen i_m berhubungan dengan jenis barang ke- m dalam transaksi pembelian.

METODOLOGI

Algoritma ini didasari bahwa untuk mencari hubungan antar *item* berdasarkan pada persepsi bahwa semakin banyak item yang dibeli dalam satu transaksi maka hubungan antar item yang terdapat dalam

transaksi ini semakin lemah. Ini untuk memperbaiki algoritma apriori yang tidak mempertimbangkan hal ini.

Data yang digunakan dalam tulisan ini diunduh dari <http://www.csd.abdn.ac.uk/hnguyen/teaching/CS553/marketbasket.csv>.

Sebelum data ini digunakan dalam algoritma yang ada, dilakukan tahap praproses data terlebih dahulu. Data asli dalam tabel transaksi berisi nilai *true* jika dibeli dan *false* jika tidak dibeli, dibuat dulu dalam data biner (0-1) artinya jika ada pembelian dinyatakan 1 dan jika tidak ada pembelian disimbolkan 0. Data 0-1 ini, kemudian dinyatakan dalam *fuzzy set* seperti pada persamaan (4). Baru kemudian dianalisa dengan algoritma yang ada. Analisis data yang dilakukan menggunakan program Microsoft Excel 2007 dan Orange 2.0b.

Langkah-langkah algoritma ini adalah :

- Menentukan *min_support* ke-*k* sebagai *threshold* bagi kombinasi *k-item* terhadap tingkat dominasinya dari keseluruhan transaksi.
- Mencari *support* dari setiap kombinasi *k-item* yang memungkinkan yang ada di dalam transaksi.
- Melakukan penyaringan terhadap kombinasi *item* yang ada di dalam transaksi tersebut yang tidak memenuhi: $support(p) \geq min_support\ ke-k$.
- Mencari kombinasi *k-item* yang memungkinkan dari tiap kombinasi (*k-1*)-*item* yang memenuhi *minimum support* yang telah ditentukan.
- Mencari *candidate rules* dengan cara menghitung *confidence* dari setiap kombinasi *k-item* yang memenuhi *min_support* ke-*k* ($k \geq 2$.)

Dalam perhitungan *support k-item*, misalkan dalam transaksi *t* terjadi pembelian barang {*a, b, d, f*}, dari kumpulan barang {*a,b,c,d,e,f*} maka nilai keenam barang tersebut dalam tabel transaksi yang akan dianalisa adalah {*a,b,c,d,e,f*} = {1/4,1/4,0,1/4,0,1/4} atau bisa dinyatakan : 1/(jumlah barang yang dibeli), jika barang

tersebut dibeli; dan 0, jika barang tersebut tidak dibeli. Ini berdasarkan persamaan (4).

Misalkan tabel transaksi di Tabel 1, diubah dalam *fuzzy set* menjadi:

Tabel 2 Tabel transaksi dalam bentuk nilai *fuzzy set*

	B1	B2	B3	B4	B5
T1	1/4	1/4	0	1/4	1/4
T2	1/4	1/4	1/4	0	1/4
T3	0	0	1	0	0
T4	1/4	1/4	0	1/4	1/4
T5	1/4	1/4	0	1/4	1/4
T6	0	1/2	0	0	1/2
T7	1/3	1/3	0	0	1/3
T8	0	0	0	1/2	1/2
T9	1/4	1/4	1/4	0	1/4
T10	0	1/3	1/3	1/3	0

Jika dibandingkan antara Tabel 1 dan Tabel 2, sebagai contoh yaitu pada Tabel 1, hubungan kedekatan *item* B4 dan B5 sebesar 1(keduanya ada) pada transaksi T1 dan T8. Tapi dalam Tabel 2 dengan memperhitungkan asumsi yang ada, maka hubungan kedekatannya *item* B4 dan B5 berbeda yaitu pada transaksi T1 sebesar 1/4 dan pada T8 sebesar 1/2.

Sekarang akan dihitung *support* untuk 1-*item set*, diperoleh : $support(B1) = 0.158$, $support(B2) = 0.242$, $support(B3) = 0.183$, $support(B4) = 0.158$, $support(B5) = 0.258$. Jika ditetapkan *min_support* sebesar 20% maka yang memenuhi hanya B2 dan B5. Selanjutnya dihitung *support* untuk 2-*item set*, diperoleh: $support(B2 \cup B5) = 0.208$.

Perhatikan bahwa nilai *support* yang diperoleh cenderung lebih kecil dibandingkan dengan menggunakan *association rule* biasa. Selain itu nilai yang diperoleh memberikan perbedaan urutan dari yang terbesar ke terkecil atau sebaliknya. Dengan menggunakan *association rule* biasa diperoleh urutan nilai *support* dari besar ke kecil adalah (B2,B5), B1, B4, B3. Sedangkan dengan algoritma ini diperoleh urutan B5, B3, (B1,B4), B2.

Selanjutnya dihitung *confidence*-nya, diperoleh: $confidence(B2 \rightarrow B5) = 0.862$ dan $confidence(B5 \rightarrow B2) = 0.806$. Jika ditetapkan

min_confidence sebesar 85 % maka *interesting rule* yang diperoleh adalah:
if B2 then B5
 [support=20.8%, confidence=86.2%],
 memberikan nilai yang cenderung lebih kecil dibandingkan dengan menggunakan *association rule* berbentuk biner dari Tabel 1.

HASIL DAN PEMBAHASAN

Berdasarkan analisis terhadap data dengan menetapkan *min_support* sebesar 1.5% untuk setiap 1-item set, dari 303 item diperoleh hanya 5 item barang yang memenuhi *min_support* yaitu barang *White Bread* (WB), *Potato Chips* (PC), *Toilet Paper* (TP), *Cream Cheese* (CC), dan *Sweet Relish* (SR), dengan nilai *support*-nya berturut-turut adalah: 0.02585, 0.02224, 0.01727, 0.01651, 0.01650.

Selanjutnya, berdasarkan barang-barang yang memenuhi *support* 1-item set dihitung *support* untuk 2-item set barang yang memenuhi *min_support* 0.1% dan diperoleh:

$$\text{support}(\text{WB} \cup \text{PC}) = 0.00336$$

$$\text{support}(\text{WB} \cup \text{TP}) = 0.00117$$

$$\text{support}(\text{PC} \cup \text{TP}) = 0.00105$$

$$\text{support}(\text{CC} \cup \text{SR}) = 0.00173$$

Penentuan *min_support* 0.1%, tidak sama dengan *min_support* 1-item set karena tidak ada yang memenuhi 1.5%. Makin mengecilnya nilai *support* karena sudah melibatkan kombinasi 2 barang yang muncul bersamaan dalam tiap transaksi. Kemudian untuk 3-item set barang, diperoleh $\text{support}(\text{WB} \cup \text{PC} \cup \text{TP}) = 0.00055$.

Selanjutnya *confidence* yang diperoleh, dengan menetapkan *min_confidence* sebesar 10%, disajikan dalam Tabel 3 berikut.

Tabel 3 Nilai *confidence* pembelian kombinasi barang

Aturan Pembelian	Nilai <i>confidence</i>
WB→PC	0.1300
PC→WB	0.1511
CC→SR	0.1047
SR→CC	0.1048
WB,PC→TP	0.1636
WB,TP→PC	0.4688
PC,TP→WB	0.5229

Berdasarkan Tabel 3 ini dapat dibuat suatu *interesting rule* yang memenuhi *min_confidence* 10 % sebagai berikut :

- Jika beli *White Bread* maka beli *Potato Chips*
 [support = 0.336%, confidence = 13%]
- Jika beli *Potato Chips* maka beli *White Bread*
 [support = 0.336%, confidence = 15.11%]
- Jika beli *Cream Cheese* maka beli *Sweet Relish*
 [support = 0.173%, confidence = 10.47%]
- Jika beli *Sweet Relish* maka beli *Cream Cheese*
 [support = 0.173%, confidence = 10.48%]
- Jika beli (*White Bread* dan *Potato Chips*) maka beli *Toilet Paper*
 [support = 0.055%, confidence = 16.36%]
- Jika beli (*White Bread* dan *Toilet Paper*) maka beli *Potato Chips*
 [support = 0.055%, confidence = 46.88%]
- Jika beli (*Potato Chips* dan *Toilet Paper*) maka beli *White Bread*
 [support = 0.055%, confidence = 52.29%].

Jika dihitung juga nilai *support* dan *confidence* menggunakan *association rule* berbentuk biner pada *k-item set* diatas, diperoleh :

$$\text{support}(\text{WB})=0.149$$

$$\text{support}(\text{PC})=0.119$$

$$\text{support}(\text{TP})=0.101$$

$$\text{support}(\text{CC})=0.83$$

$$\text{support}(\text{SR})=0.95$$

$$\text{support}(\text{WB} \cup \text{PC})=0.71$$

$$\text{support}(\text{WB} \cup \text{TP})=0.48$$

$$\text{support}(\text{PC} \cup \text{TP})=0.54$$

$$\text{support}(\text{CC} \cup \text{SR})=0.36$$

$$\text{support}(\text{WB} \cup \text{PC} \cup \text{TP})=0.36$$

$confidence(WB \rightarrow PC) = 0.48$
 $confidence(PC \rightarrow WB) = 0.60$
 $confidence(CC \rightarrow SR) = 0.429$
 $confidence(SR \rightarrow CC) = 0.375$
 $confidence((WB \cup PC) \rightarrow TP) = 0.5$
 $confidence((WB \cup TP) \rightarrow PC) = 0.75$
 $confidence((PC \cup TP) \rightarrow WB) = 0.667$.

Nilai *support* dan *confidence* yang diperoleh cenderung lebih kecil namun lebih realistis dibandingkan yang menggunakan *association rule* biasa, dikarenakan dalam perhitungannya menggunakan nilai *fuzzy set* yang menyatakan tingkat asosiasi antara *item* barang makin kecil dengan makin banyaknya barang yang dibeli dalam satu transaksi.

KESIMPULAN

Berdasarkan analisa dalam tulisan ini ternyata tabel transaksi bisa dinyatakan dalam bentuk *fuzzy set* yang mengasumsikan bahwa makin banyak barang yang dibeli dalam satu transaksi maka tingkat asosiasinya makin kecil.

Dibandingkan dengan menggunakan *association rule* berbentuk biner, dengan menggunakan *fuzzy c-partition* memberikan nilai *support* dan *confidence* cenderung lebih kecil tapi ini memberikan hasil yang lebih realistis.

DAFTAR PUSTAKA

- Auslender L.E. 2004. *On Analytical Tools for Market Basket Analysis*. SAS Institute, Research & Development. New York.
- Han J., M. Kamber. 2006. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publisher.
- Helm L. 2007. *Fuzzy Association Rules. An Implementation in R* [Thesis]. Vienna University. Vienna.
- <http://www.csd.abdn.ac.uk/~hnguyen/teaching/CS5553/marketbasket.csv>.
- Intan R. 2002. *Degree of Similarity in Fuzzy Partition*. Lecture Notes in Computer Science 2275:99-107.
- Intan R. 2006. A Proposal of Multidimensional Association Rules. *Jurnal Informatika* 7(2): 85-90.
- Intan R., M. Masao. 2003. A Proposal of Fuzzy Thesaurus Generated by Fuzzy Covering. *IEEE*. 2003: 167-172.
- Klir G. J., B.Yuan. 2001. *Fuzzy Sets and Fuzzy Logic-Theory and Applications*. Prentice Hall India.
- Possas B., M. Wagner, C. Marcio, R. Rodolfo, 2000. Using Quantitative Information for Efficient Association Rule Generation, *ACM SIGMOD* 29:19-25.
<http://csis.pace.edu/csis/masplas/p8.pdf>
- Tan P.N., M. Steinbach, V. Kumar. 2006. *Introduction to Data Mining*. Addison Wesley. Boston.
- Ulmer D. 2002. Mining an Online Auctions Data Warehouse, *Proceedings of MASPLAS'02 The Mid-Atlantic Student Workshop on Programming Languages and Systems, Pace University*.