

# IMPLEMENTASI ALGORITMA APRIORI UNTUK MENEMUKAN *FREQUENT ITEMSET* DALAM KERANJANG BELANJA

Adie Wahyudi Oktavia Gama<sup>1</sup>, I Ketut Gede Darma Putra<sup>2</sup>, I Putu Agung Bayupati<sup>3</sup>

**Abstract**— Apriori algorithm use an iterative approach where  $k$ -itemset used to explore  $(k+1)$ -itemset.  $(K+1)$ -itemset candidates containing subset frequency that rarely appears will not used in determining association rules. Association rules formed by “if antecedent then consequent”. Implementation of the apriori algorithm was preceded by the preparation of transactions database and determination of minimum support and confidence. Apriori algorithm scanning database repeated, pair one item to another and record the number of occurrences in the overall transaction. Frequent itemset is determined by selecting a combination or itemset that the count value greater than or equal to the minimum support then calculated the percentage value of support and confidence of each candidate. The association rules selected from which fit the minimum support and confidence. Data used in this study was sample of 100 transactions from database point of sales. Final Association rules are obtained from implementation of apriori algorithm is “if HELLO PANDA REFIL then HELLO PANDA 10gr” with percentage of support 2.00% and confidence 100.00% and “if HELLO PANDA 10gr then HELLO PANDA REFIL” with percentage of support 2.00% and confidence 100.00%. So it can be concluded that most customers buy HELLO PANDA REFIL will also buy HELLO PANDA 10gr so goes otherwise. This study proves that the apriori algorithm suitable implemented to search the frequent itemset in the shopping cart. Association rules that formed of frequent itemset can be used as decision support in sales.

**Keywords**— Data Mining, Association Rules, Apriori Algorithm, Frequent Itemset, Item Combination, Shopping Cart.

**Intisari**— Algoritma apriori menggunakan pendekatan iteratif dimana  $k$ -itemset digunakan untuk mengeksplorasi  $(k+1)$ -itemset. Calon  $(k+1)$ -itemset yang mengandung frekuensi subset yang jarang muncul tidak dipakai menentukan aturan asosiasi. Aturan asosiasi berbentuk *if antecedent then consequent*. Implementasi algoritma apriori didahului dengan persiapan database transaksi serta penentuan batas minimum support dan confidence. Algoritma apriori melakukan scanning database berulang-ulang, memasangkan satu item dengan item lainnya dan mencatat jumlah kemunculan kombinasi dalam keseluruhan transaksi. Frequent itemset ditentukan dengan memilih itemset yang nilai kemunculannya diatas atau sama dengan nilai minimum support kemudian dihitung persentase nilai support dan confidencenya. Aturan asosiasi yang berlaku dipilih dari yang memenuhi syarat minimum support dan confidence. Data yang digunakan dalam penelitian ini adalah sampel 100 transaksi dari database point of sales. Aturan asosiasi final yang diperoleh dari penerapan algoritma apriori adalah “if HELLO PANDA REFIL then HELLO PANDA 10GR” dengan persentase support 2,00% dan confidence 100,00% dan aturan “if HELLO PANDA 10GR then HELLO PANDA REFIL” dengan persentase support 2,00% dan confidence 100,00%. Jadi dapat disimpulkan kebanyakan pelanggan membeli item HELLO PANDA REFIL juga akan membeli item HELLO PANDA 10GR begitu juga berlaku

sebaliknya. Penelitian ini membuktikan bahwa algoritma apriori cocok diimplementasikan untuk mencari frequent itemset pada keranjang belanja. Aturan asosiasi yang dibentuk dari frequent itemset tersebut dapat dipakai sebagai pendukung keputusan dalam penjualan.

**Kata Kunci**— Data Mining, Association Rules, Apriori Algorithm, Frequent Itemset, Item Combination, keranjang belanja

## I. PENDAHULUAN

Data mining adalah proses untuk mengekstraksi atau menggali pengetahuan dari data yang berjumlah besar [1]. Data berskala besar memunculkan fenomena “data rich but information poor”, dimana data yang berjumlah besar tidak diikuti dengan peningkatan informasi yang didapat dari data tersebut. Fenomena itu sering muncul pada sistem penjualan yang dipakai di pasar swalayan, dimana data transaksi terus bertambah setiap hari tetapi hanya menghasilkan laporan transaksi harian yang monoton. Hal ini dikarenakan kurangnya analisa terhadap kumpulan data tersebut.

Analisis asosiasi adalah salah satu teknik data mining yang digunakan untuk menemukan hubungan menarik antara suatu kombinasi item yang tersembunyi dalam suatu database. Hubungan ini dapat direpresentasikan dalam suatu bentuk aturan asosiasi [2]. Analisis asosiasi akan berusaha mengungkap asosiasi antara atribut, yaitu berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut. Secara umum aturan asosiasi mempunyai bentuk :

Kekuatan hubungan suatu aturan asosiatif dapat diukur dengan dua parameter yaitu support dan confidence. Support (nilai penunjang) adalah persentase kombinasi item tersebut dalam database dan confidence (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiatif yang terbentuk oleh metode asosiasi dalam data mining.

Algoritma apriori termasuk jenis aturan asosiasi pada data mining yang didasarkan pada fakta bahwa apriori menggunakan pengetahuan sebelumnya dari suatu itemset dengan frekuensi kemunculan yang sering atau disebut dengan frequent itemset. Apriori menggunakan pendekatan iteratif dimana  $k$ -itemset digunakan untuk mengeksplorasi  $(k+1)$ -itemset [1]. Pada algoritma ini calon  $(k+1)$ -itemset dihasilkan oleh penggabungan dua itemset pada domain/ukuran  $k$ . Calon  $(k+1)$ -itemset yang mengandung frekuensi subset yang jarang muncul atau dibawah threshold akan dipangkas dan tidak dipakai dalam menentukan aturan asosiasi [2].

Dengan latar belakang masalah tersebut maka dalam penelitian ini akan membahas tentang implementasi algoritma apriori dalam keranjang belanja untuk menemukan frequent itemset.

<sup>1</sup>Mahasiswa Magister Teknik Elektro, Universitas Udayana, Kampus Universitas Udayana Jl. P.B. Sudirman Denpasar Bali 80232 (Telp/Fax: (0361) 239599; e-mail: aade.1025@gmail.com)

<sup>2,3</sup>Dosen Magister Teknik Elektro, Universitas Udayana, Kampus Universitas Udayana Jl. P.B. Sudirman Denpasar Bali 80232 (Telp/Fax: (0361) 239599; e-mail: <sup>2</sup>ikgdarmaputra@gmail.com,

<sup>3</sup>bayuhelix@yahoo.com)



## II. STUDI LITERATUR

### A. Metode Asosiasi

Metode analisis asosiasi, juga dikenal sebagai *market basket analysis*, yaitu analisis yang sering dipakai untuk menganalisa isi keranjang belanja konsumen dalam suatu pasar swalayan. Contoh penerapan dari aturan asosiatif adalah analisa pembelian produk pada sebuah toko alat tulis, pada analisa itu misalkan dapat diketahui berapa besar kemungkinan seorang pelanggan membeli pensil bersamaan dengan membeli penghapus. Penerapan aturan asosiasi dalam kasus tersebut dapat membantu pemilik toko untuk mengatur penempatan barang, mengatur persediaan atau membuat promosi pemasaran dengan menerapkan diskon untuk kombinasi barang tertentu.

Analisis asosiasi didefinisikan sebagai suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk *support* (*minimum support*) dan syarat minimum untuk *confidence* (*minimum confidence*). Dasar analisis asosiasi terbagi menjadi dua tahap, yaitu:

1. Analisa pola frekuensi tinggi, pada tahap ini dicari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus berikut :

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

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

$$\text{Support}(A, B) = P(A \cap B) \quad (2)$$

$$\text{Support}(A, B) = \frac{\sum \text{Transaksi untuk } A \text{ dan } B}{\sum \text{Transaksi}} \quad (3)$$

2. Pembentukan aturan Asosiasi, setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung nilai *confidence* aturan asosiatif  $A \rightarrow B$ .

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

$$\text{Confidence} = P(B|A) = \frac{\sum \text{Transaksi untuk } A \text{ dan } B}{\sum \text{Transaksi untuk } A} \quad (4)$$

### B. Algoritma Apriori

Algoritma apriori termasuk jenis aturan asosiasi pada *data mining*. Penelitian yang dilakukan R. Agrawal, et al pada tahun 1993 yang berjudul "*Mining Association Rules Between Sets of Items in Large Databases*" adalah awal mula dikembangkannya *data mining* dengan metode asosiasi menggunakan algoritma apriori [3]. Pada tahun 1994, R. Agrawal dan R. Srikant kembali melakukan penelitian mengenai metode asosiasi dengan judul "*Fast Algorithms for Mining Association Rules*". Penelitian ini kemudian difokuskan untuk menyempurnakan algoritma apriori yang sudah dikembangkan sebelumnya dan dari situlah algoritma apriori dikenal sebagai salah satu algoritma untuk metode asosiasi [4]. Apriori menggunakan pendekatan iteratif dimana *k-itemset* digunakan untuk mengeksplorasi  $(k+1)$ -*itemset*

berikutnya [1]. Prinsip metode apriori adalah jika suatu *itemset* sering muncul (*frequent*), maka semua *subset* dari *itemset* tersebut juga harus sering muncul dalam suatu *database* [2].

Pada algoritma ini calon  $(k+1)$ -*itemset* dihasilkan oleh penggabungan dua *itemset* pada domain / ukuran  $k$ . Calon  $(k+1)$ -*itemset* yang mengandung frekuensi *subset* yang jarang muncul atau dibawah *threshold* akan dipangkas dan tidak dipakai dalam menentukan aturan asosiasi [2]. Sesuai dengan aturan asosiasi, algoritma apriori juga menggunakan *minimum support* dan *minimum confidence* untuk menentukan aturan *itemset* mana yang sesuai untuk digunakan dalam pengambilan keputusan.

1-*itemset* digunakan untuk menemukan 2-*itemset* yaitu kombinasi *item* yang berjumlah 2, contohnya *if buy pensil then buy penghapus*, 2-*itemset* digunakan untuk menemukan 3-*itemset* yaitu kombinasi *item* yang berjumlah 3, contohnya *if buy pensil and buy pulpen then buy penghapus* dan begitu seterusnya sampai tidak ada lagi *frequent k-itemset* yang bisa ditemukan dalam *database* transaksi [1].

Struktur dari *itemset* disini adalah mengikuti suatu bentuk dari kombinasi. Pengertian kombinasi adalah menggabungkan beberapa objek dari suatu grup tanpa memperhatikan urutan [5]. Di dalam kombinasi, urutan objek tidak diperhatikan sebagai contoh dimana {1,2,3} adalah sama dengan {2,3,1} dan {3,1,2}.

Kombinasi dapat dibagi menjadi dua yaitu kombinasi dengan pengulangan dan kombinasi tanpa pengulangan. Kombinasi yang dipakai dalam algoritma apriori dalam penelitian ini adalah kombinasi tanpa pengulangan karena urutan tidak diperhatikan akan tetapi setiap *item* yang ada hanya digunakan sekali dalam satu kombinasi *itemset*.

Kombinasi tanpa pengulangan ketika urutan tidak diperhatikan akan tetapi setiap objek yang ada hanya bisa dipilih sekali maka jumlah kombinasi yang ada adalah:

$$\frac{n!}{r!(n-r)!} = \binom{n}{r} \quad (5)$$

Dimana  $n$  adalah jumlah objek yang bisa dipilih dan  $r$  adalah jumlah yang harus dipilih.

Sebagai contoh, terdapat 5 pensil warna dengan warna yang berbeda yaitu; merah, kuning, hijau, biru dan ungu. Pensil warna tersebut hanya boleh dipilih dua warna. Banyak cara untuk mengkombinasikan pensil warna yang ada dengan menggunakan rumus di atas adalah  $5!/(5-2)!(2)! = 10$  kombinasi.

Penelitian tentang algoritma apriori sudah banyak dilakukan sebelumnya. Penelitian-penelitian lain yang berhubungan dengan algoritma apriori yang dipakai sebagai referensi dalam penelitian ini antara lain adalah sebagai berikut:

1. Jogi Suresh dan T. Ramanjaneyulu (2013) melakukan penelitian dengan judul "*Mining Frequent Itemsets Using Apriori Algorithm*". Penelitian Suresh dan Ramanjaneyulu menggunakan algoritma apriori klasik yang sudah dikembangkan sebelumnya dan belum menggunakan teknik optimasi untuk memperoleh aturan asosiasi yang lebih efisien [6].
2. Sheila A. Abaya pada tahun 2012 dalam penelitiannya yang berjudul "*Association Rule Mining based on Apriori*

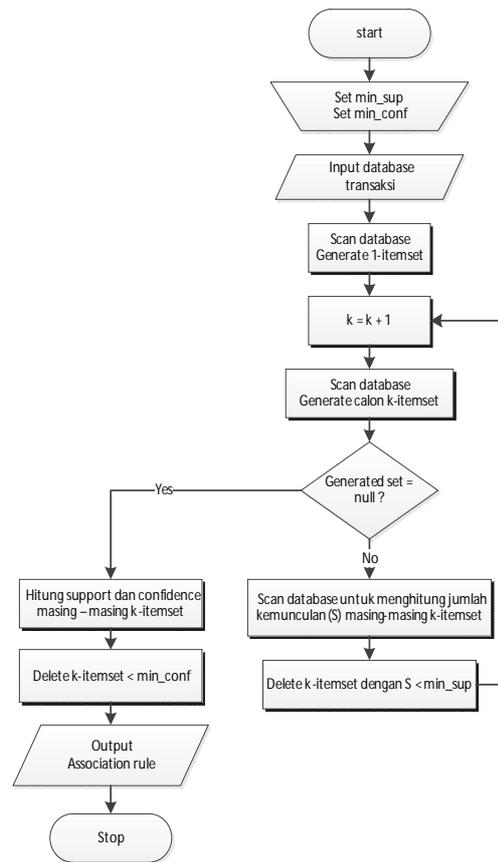
Algorithm in Minimizing Candidate Generation” melakukan improvisasi terhadap algoritma apriori. Improvisasi dilakukan dengan cara menentukan “set size” dan “set size frequency”. Set size adalah jumlah item per transaksi sedangkan set size frequency adalah jumlah transaksi yang setidaknya memiliki “set size” item. Set size dan set size frequency ini digunakan untuk mengeliminasi kandidat kunci yang tidak signifikan [7].

3. Jiao Yabing (2013) dalam penelitiannya dengan judul “Research of an Improved Apriori Algorithm in Data mining Association Rules” melakukan optimasi terhadap algoritma apriori yaitu dengan cara mengurangi atau memangkas (pruning) jumlah calon kandidat frequent itemset pada kandidat itemset Ck [8].
4. Jaishree Singh, et al pada tahun (2013) melakukan penelitian dengan judul “Improving Efficiency of Apriori Algorithm Using Transaction Reduction”. Penelitian Singh, dkk ini melakukan improvisasi algoritma apriori dengan cara mengurangi jumlah transaksi (transaction reduction) yang jumlah item pertransaksinya tidak memenuhi nilai batas yang ditentukan. Pengurangan transaksi tersebut berdampak pada efisiensi waktu yang lebih cepat saat scanning database [9].

### III. METODE

Algoritma apriori digunakan dalam penelitian ini adalah untuk menemukan frequent itemset yang tersembunyi pada database transaksi atau keranjang belanja. Frequent itemset tersebut kemudian digunakan untuk membuat aturan asosiasi dengan syarat memenuhi batas minimum support dan confidence. Flowchart atau bagan alir dari algoritma apriori yang diimplementasikan pada database transaksi untuk menemukan frequent itemset ditunjukkan pada Gambar 1.

1. Apriori menggunakan pendekatan iteratif dimana k-itemset digunakan untuk mengeksplorasi (k+1)-itemset.
2. Calon (k+1)-itemset dengan frekuensi yang jarang muncul atau dibawah threshold (min\_support) akan dipangkas dan tidak dipakai dalam menentukan aturan asosiasi.
3. Langkah pertama adalah 1-itemset ditemukan dengan melakukan scanning database untuk mengakumulasi jumlah dari setiap kategori item dan kemunculannya pada setiap transaksi.
4. Selanjutnya 1-itemset digunakan untuk menemukan 2-itemset. Calon kandidat 2-itemset ditemukan dengan cara saling memasangkan satu item dengan item yang lainnya sehingga didapat kombinasi yang memungkinkan untuk 2 buah item.
5. 2-itemset kemudian dihitung nilai kemunculannya pada setiap transaksi. Nilai ambang (min\_support) ditentukan untuk memangkas calon kandidat yang dianggap tidak sering muncul atau tidak frequent.
6. 2-itemset yang memenuhi nilai ambang akan dihitung nilai support dan confidence-nya. 2-itemset yang memenuhi min\_support dan min\_confidence akan dijadikan aturan asosiasi.
7. 2-itemset digunakan untuk menemukan 3-itemset dan seterusnya sampai tidak ada lagi frequent (k+1)-itemset yang bisa ditemukan



Gambar 1: Flowchart Algoritma Apriori

8. Setelah semua aturan dari frequent (k+1)-itemset terbentuk, kemudian dihitung nilai support dan confidence. Hasil perkalian dari nilai support dan confidence yang paling tinggi merupakan aturan asosiasi yang paling baik dari keseluruhan transaksi yang ada dalam database.

#### A. Penerapan Algoritma Apriori pada Data Transaksi Penjualan

Contoh berikut diberikan lima buah data transaksi penjualan barang seperti dijabarkan pada tabel berikut:

Tabel 1: Data Transaksi Penjualan

TId	Tanggal Transaksi	Nama Barang
1	2013-06-10	Bolpoin, Buku Tulis
2	2013-06-10	Buku Tulis, Pensil, Penghapus
3	2013-06-10	Pensil, Penghapus
4	2013-06-10	Pensil, Buku Gambar, Penghapus
5	2013-06-10	Pensil, Penghapus, Bolpoin, Buku Tulis

Data transaksi tersebut dijabarkan dalam bentuk tabular menjadi 1-itemset seperti tabel dibawah ini. Fungsinya nanti adalah untuk mendapatkan calon (k+1)-itemset berikutnya:



Tabel 2: Penjabaran Transaksi dalam Bentuk 1-Itemset

No	Bolpoin	Buku Tulis	Pensil	Penghapus	Buku Gambar
1	1	1	0	0	0
2	0	1	1	1	0
3	0	0	1	1	0
4	0	0	1	1	1
5	1	1	1	1	0
Jumlah	2	3	4	4	1

Buat kombinasi 2-itemset yang mungkin, dengan memasangkan satu item dengan item lainnya. Kombinasi yang telah dibentuk kemudian hitung jumlah kemunculannya pada setiap transaksi.

Tabel 3: Calon Kandidat 2-Itemset yang mungkin

Kombinasi	Jumlah
Bolpoin, Buku Tulis	2
Bolpoin, Pensil	1
Bolpoin, Penghapus	1
Bolpoin, Buku Gambar	0
Buku Tulis, Pensil	2
Buku Tulis, Penghapus	2
Buku Tulis, Buku Gambar	0
Pensil, Penghapus	4
Pensil, Buku Gambar	1
Penghapus, Buku Gambar	1

Jika ditetapkan nilai *threshold* (*min\_support*) = 2, maka didapat *frequent 2-itemset* (F2) yaitu :

F2 = [Bolpoin, Buku Tulis], [Buku Tulis, Pensil], [Buku Tulis, Penghapus], [Pensil, Penghapus]

Calon kandidat 3-itemset dibentuk juga dengan cara yang sama yaitu memasangkan item satu dengan item lain sehingga membentuk calon kandidat 3-itemset seperti pada Table 4.

Tabel 4: Calon Kandidat 3-Itemset yang mungkin

Kombinasi	Jumlah
Bolpoin, Buku Tulis, Pensil	1
Bolpoin, Buku Tulis, Penghapus	1
Bolpoin, Buku Tulis, Buku Gambar	0
Bolpoin, Pensil, Penghapus	1
Bolpoin, Pensil, Buku Gambar	0
Bolpoin, Penghapus, Buku Gambar	0
Buku Tulis, Pensil, Penghapus	2
Buku Tulis, Pensil, Buku Gambar	0
Buku Tulis, Penghapus, Buku Gambar	0
Pensil, Penghapus, Buku Gambar	1

Nilai *threshold* (*min\_support*) sudah ditetapkan sebelumnya = 2, maka didapat *frequent 3-itemset* (F3) yaitu:

F3 = [Buku Tulis, Pensil, Penghapus]

Jika tidak terdapat lagi (k+1)-itemset yang bisa dibentuk, maka hitung nilai *support* dan *confidence* masing-masing *frequent itemset*. Aturan asosiasi dibentuk berdasarkan *frequent (k+1)-itemset* yang terpilih.

Tabel 5: Calon Aturan Asosiasi dari Frequent 2-Itemset (F2)

IF Antecedent, then Consequent	Support	Confidence
If Bolpoin, then Buku Tulis	2 / 5 = 40%	2 / 2 = 100%
If Buku Tulis, then Bolpoin	2 / 5 = 40%	2 / 3 = 66.7%
If Buku Tulis, then Pensil	2 / 5 = 40%	2 / 3 = 66.7%
If Pensil, then Buku Tulis	2 / 5 = 40%	2 / 4 = 50%
If Buku Tulis, then Penghapus	2 / 5 = 40%	2 / 3 = 66.7%
If Penghapus, then Buku Tulis	2 / 5 = 40%	2 / 4 = 50%
If Pensil, then Penghapus	4 / 5 = 80%	4 / 4 = 100%
If Penghapus, then Pensil	4 / 5 = 80%	4 / 4 = 100%

Tabel 6: Calon Aturan Asosiasi dari Frequent 3-Itemset (F3)

IF Antecedent, then Consequent	Support	Confidence
If Buku Tulis and Pensil, then Penghapus	2 / 5 = 40%	2 / 2 = 100%
If Buku Tulis and Penghapus, then Pensil	2 / 5 = 40%	2 / 2 = 100%
If Pensil and Penghapus, then Buku Tulis	2 / 5 = 40%	2 / 4 = 50%

Aturan asosiasi yang dipilih adalah aturan yang mempunyai nilai *confidence* yang lebih besar atau sama dengan dari nilai *min\_confidence*. Ditetapkan nilai *min\_confidence* adalah 80%, maka aturan asosiasi final yang dapat dibentuk adalah seperti pada Table 7:

Tabel 7: Aturan Asosiasi Final

If Antecedent, then Consequent	Support	Confidence	Support * Confidence
if buy Bolpoin, then buy Buku Tulis	2 / 5 = 40%	2 / 2 = 100%	0.4
if buy Pensil, then buy Penghapus	4 / 5 = 80%	4 / 4 = 100%	0.8
if buy Penghapus, then buy Pensil	4 / 5 = 80%	4 / 4 = 100%	0.8
if buy Buku Tulis and Pensil, then buy Penghapus	2 / 5 = 40%	2 / 2 = 100%	0.4
if buy Buku Tulis and Penghapus, then buy Pensil	2 / 5 = 40%	2 / 2 = 100%	0.4

B. Aturan Asosiasi

Aturan asosiasi merupakan hasil akhir yang ingin dicapai yaitu bertujuan untuk memilih aturan yang paling cocok untuk dapat digunakan sebagai pedoman pengambilan keputusan dan strategi pemasaran yang lebih baik. Tahap ini menghasilkan output berupa *frequent itemset* atau *rule* dengan nilai perkalian *support* dan *confidence* yang paling tinggi. Tahap ini merupakan kesimpulan akhir dari proses apriori yang nantinya menjelaskan bahwa aturan asosiasi yang mempunyai pengaruh paling kuat adalah aturan yang memiliki nilai perkalian *support* dan *confidence* yang paling tinggi.

IV. HASIL DAN PEMBAHASAN

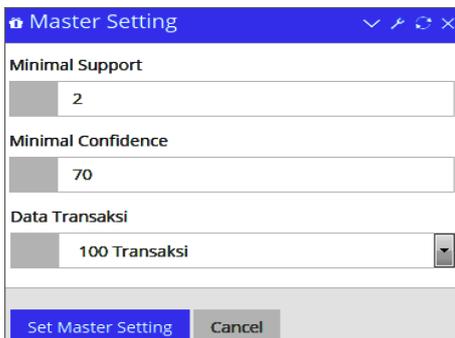
A. Identifikasi Awal

Identifikasi awal berguna untuk memberikan batasan-batasan pada algoritma apriori agar berjalan sesuai dengan apa yang telah ditentukan. Identifikasi awal tersebut termasuk didalamnya adalah menentukan batas *minimum support* dan

minimum confidence yang digunakan dalam membuat aturan asosiasi serta database transaksi yang digunakan untuk menguji algoritma apriori. Minimum support dan minimum confidence adalah nilai batas (threshold) yang nantinya digunakan untuk menentukan aturan asosiasi yang terbaik.

1) Menentukan Batas Minimum Support dan Minimum Confidence

Apriori menggunakan pendekatan iteratif dimana k-itemset digunakan untuk mengeksplorasi (k+1)-itemset [1]. Calon (k+1)-itemset dengan frekuensi yang jarang muncul atau dibawah threshold / minimum support akan dipangkas dan tidak dipakai dalam menentukan (k+1)-itemset selanjutnya. Setelah tidak ditemukan lagi kombinasi (k+1)-itemset yang mungkin maka dibentuk beberapa rule atau aturan dengan kombinasi antar-item yang paling kuat yang sering muncul dalam proses transaksi diatas ambang minimum confidence disebut dengan aturan asosiasi.



Gambar 2: Tampilan Menentukan Nilai Minimum Support dan Minimum Confidence

Nilai minimum support dan minimum confidence ditentukan melalui proses pengamatan dari beberapa kali percobaan pada database yang digunakan. Penentuan nilai tersebut disesuaikan agar menghasilkan aturan asosiasi yang baik. Nilai minimal support yang ditentukan dalam penelitian ini adalah 2 karena minimal jumlah item untuk membentuk suatu kombinasi adalah 2 item. Alasan lain adalah karena dari beberapa kali percobaan jika ditentukan nilai minimum support lebih dari itu dari data transaksi yang tersedia maka tidak ditemukan aturan asosiasi yang baik karena jumlah kemunculan suatu kombinasi paling sering adalah 2 kali dalam keseluruhan transaksi. Nilai minimum confidence ditentukan sebesar 70% untuk menghasilkan aturan yang terbaik. Kombinasi aturan yang memenuhi nilai tersebut berarti mempunyai nilai hubungan keterkaitan antar item yang lebih kuat dengan kemungkinan kemunculan sebanyak 70% dari suatu item dengan item pasangannya dalam satu itemset.

2) Database Transaksi

Data yang dipakai dalam penelitian ini adalah berasal dari sampel database sistem point of sales Toko Tujuh Belas. Data yang tersimpan dalam database adalah data transaksi penjualan dari bulan Juni 2011 sampai dengan bulan Oktober 2012. Sampel data yang dipakai adalah 100 transaksi yang sudah melalui tahap preprocessing dengan membersihkan data yang rusak.

B. Pembentukan Kombinasi Calon (k+1) Itemset

Pembentukan kombinasi didapat dari proses scanning database yang berulang-ulang dengan memasangkan satu item dengan item lainnya. Kombinasi yang dibentuk adalah kombinasi tanpa pengulangan ketika urutan tidak diperhatikan akan tetapi setiap item yang ada hanya bisa dipilih sekali dalam membentuk kombinasi. Berikut adalah beberapa calon kombinasi yang dibangkitkan dari item-item dalam transaksi.

- Calon kombinasi dengan 2-itemset

Calon 2-itemset		
Kombinasi		Jumlah
NARIS1500	INDOMIE KARI AYAM 17GR	0
NARIS1500	NARSIS 2500	0
NARIS1500	SPECIAL CREAM COKLAT	0
NARIS1500	TELUR PER BUTIR	0
NARIS1500	KUE BASAH/ROTI CAMPUR	1
NARIS1500	PIA CEMERLANG ISI 9	0
NARIS1500	KACANG BOBY	0
NARIS1500	MARLBORO PUTIH	1

Gambar 3: Calon Kombinasi dengan 2-itemset

- Calon kombinasi dengan 3-itemset

Calon 3-itemset			
Kombinasi			Jumlah
NARIS1500	INDOMIE KARI AYAM 17GR	NARSIS 2500	0
NARIS1500	INDOMIE KARI AYAM 17GR	SPECIAL CREAM COKLAT	0
NARIS1500	INDOMIE KARI AYAM 17GR	TELUR PER BUTIR	0
NARIS1500	INDOMIE KARI AYAM 17GR	KUE BASAH/ROTI CAMPUR	0
NARIS1500	INDOMIE KARI AYAM 17GR	PIA CEMERLANG ISI 9	0
NARIS1500	INDOMIE KARI AYAM 17GR	KACANG BOBY	0
NARIS1500	INDOMIE KARI AYAM 17GR	MARLBORO PUTIH	0

Gambar 4: Calon Kombinasi dengan 3-itemset

Calon kombinasi akan terus dicari sampai sebanyak (k+1)-itemset. Iterasi akan berhenti sampai tidak ditemukannya lagi kombinasi atau itemset yang mungkin pada database transaksi untuk dapat membentuk itemset berikutnya

C. Menentukan Frequent (K+1) Itemset

Kombinasi-kombinasi yang telah dibentuk kemudian dihitung nilai kemunculannya dalam keseluruhan transaksi. Nilai kemunculan kombinasi dihitung dengan cara scanning database berulang-ulang dan mencatat berapa kali kombinasi tersebut muncul dalam keseluruhan transaksi.

Kombinasi-kombinasi tersebut kemudian dipilih mana yang termasuk frequent itemset dengan cara memilih pasangan atau itemset yang nilai kemunculannya diatas atau sama dengan nilai minimum support yang telah ditentukan dalam keseluruhan transaksi.



Dari data tersebut diatas, jika ditetapkan nilai min kemunculan = 2 maka, F2 :

Kombinasi		Jumlah
HELLO PANDA REFIL	HELLO PANDA 10GR	2

Gambar 5. Itemset dengan Nilai Kemunculan di Atas atau Sama Dengan Minimum Support

Hasil diatas menunjukkan bahwa pasangan kombinasi HELLO PANDA REFIL dan HELLO PANDA 10GR adalah pasangan yang memenuhi syarat sebagai *frequent itemset* dengan nilai kemunculan 2 kali dalam 100 transaksi, sedangkan kombinasi-kombinasi yang lain kemunculannya kurang dari 2 kali dalam 100 transaksi. Kombinasi tersebut kemudian dibuatkan calon aturan asosiasi dan dihitung nilai *support* dan *confidencinya*.

#### D. Pembentukan Aturan Asosiasi

Aturan asosiasi dalam aplikasi ini dipilih dari kombinasi *frequent itemset* yang memenuhi nilai *minimum support* yang sudah ditentukan pada proses sebelumnya. Calon aturan asosiasi dibuatkan dari kombinasi tersebut dan dihitung persentase *support* dan *confidence* dari masing-masing kombinasi. *Item-item* dalam kombinasi bergiliran menjadi *antecedent* dan *consequent*. Jika satu *item* menjadi *antecedent* maka *item-item* yang lain menjadi *consequent*.

Calon aturan dari kombinasi *frequent itemset* HELLO PANDA REFIL DAN HELLO PANDA 10GR perhitungan *support* dan *confidencinya* adalah sesuai dengan tabel berikut:

Tabel 8: Calon Aturan Asosiasi dari *Frequent 2-Itemset* (F2)

IF Antecedent, then Consequent	Support	Confidence
If HELLO PANDA REFIL, then HELLO PANDA 10GR	$2 / 100 = 2.00\%$	$2 / 2 = 100.00\%$
If HELLO PANDA 10GR, then HELLO PANDA REFIL	$2 / 100 = 2.00\%$	$2 / 2 = 100.00\%$

Perhitungan nilai *support* dan *confidence* didapat dari dasar analisis asosiasi yang dibagi menjadi dua tahap sesuai dengan teori yang sudah dipaparkan sebelumnya, yaitu seperti Gambar 6.:

Calon aturan yang memenuhi nilai minimal *confidence* 70% adalah *antecedent* dan *consequent* dengan bentuk aturan *if HELLO PANDA REFIL then HELLO PANDA 10GR* yang mempunyai persentase nilai *support* 2,00% dan persentase nilai *confidence* 100,00% dan juga aturan *if HELLO PANDA 10GR then HELLO PANDA REFIL* yang juga mempunyai persentase nilai *support* 2,00% dan persentase nilai *confidence* 100,00%. Jadi dapat disimpulkan kebanyakan pelanggan membeli HELLO PANDA REFIL juga akan membeli HELLO PANDA 10GR begitu juga berlaku sebaliknya.

Berdasarkan analisis dan pembahasan tersebut diatas maka dapat dilihat bahwa algoritma apriori dapat menemukan *frequent itemset* tersembunyi dari keranjang belanja yang digunakan untuk membentuk aturan asosiasi. Aturan asosiasi yang dihasilkan kemudian digunakan sebagai pendukung keputusan yang baik dalam penjualan.

Kombinasi		Jumlah	
HELLO PANDA REFIL	HELLO PANDA 10GR	2	

Calon Aturan Asosiasi dari F2

IF Antecedent, then Consequent	Support	Confidence
if HELLO PANDA REFIL then HELLO PANDA 10GR	2.00%	100.00%
if HELLO PANDA 10GR then HELLO PANDA REFIL	2.00%	100.00%

ATURAN ASSOSIASI FINAL

IF Antecedent, then Consequent	Support	Confidence	Support * Confidence
if HELLO PANDA REFIL then HELLO PANDA 10GR	2.00%	100.00%	200.00%
if HELLO PANDA 10GR then HELLO PANDA REFIL	2.00%	100.00%	200.00%

Gambar 6: Pembentukan Aturan Asosiasi

## V. KESIMPULAN

Data yang digunakan dalam penelitian ini adalah sampel 100 transaksi dari *database* sistem *point of sales*. Aturan asosiasi final yang diperoleh dari penerapan algoritma apriori pada sampel database tersebut adalah “*if HELLO PANDA REFIL then HELLO PANDA 10GR*” dengan persentase *support* 2,00% dan *confidence* 100,00% dan juga aturan “*if HELLO PANDA 10GR then HELLO PANDA REFIL*” dengan persentase *support* 2,00% dan *confidence* 100,00%. Jadi dapat disimpulkan kebanyakan pelanggan membeli *item* HELLO PANDA REFIL juga akan membeli *item* HELLO PANDA 10GR begitu juga berlaku sebaliknya. Metode asosiasi menggunakan algoritma apriori cocok diimplementasikan untuk menemukan *frequent itemset* yang tersembunyi pada keranjang belanja. Aturan asosiasi yang dihasilkan dari *frequent itemset* dapat dipakai sebagai pendukung keputusan dalam penjualan seperti mengatur penempatan barang, mengatur persediaan atau membuat promosi pemasaran dengan menerapkan diskon untuk kombinasi barang-barang tertentu yang sering muncul dalam transaksi.

## REFERENSI

- [1]. J. Han and M. Kamber, *Data Mining Concept and Techniques*, 2nd edition, USA: Elsevier, Inc, 2006.
- [2]. P. N. Tan, et al, *Introduction to Data Mining*, USA: Addison-Wesley, 2004.
- [3]. R. Agrawal, et al, *Mining Association Rules between Sets of Items in Large Databases*. Proceedings of the 1993 ACM SIGMOD Conference Washington DC, USA. 1993.
- [4]. R. Agrawal and R. Srikant, *Fast Algorithm for Mining Association Rules*. Proceeding of the 20<sup>th</sup> VLDB Conference Santiago, Chile. 1994.
- [5]. Wikipedia. *Kombinasi dan Permutasi*. Available at [https://id.wikipedia.org/wiki/Kombinasi\\_dan\\_permutasi](https://id.wikipedia.org/wiki/Kombinasi_dan_permutasi). Acces date on January 5<sup>th</sup>, 2016. 2016.
- [6]. J. Suresh, T. Ramanjaneyulu, “Mining Frequent Itemsets Using Apriori Algorithm,” *International Journal of Computer Trends and Technology*, 2013 (tidak ada vol, no dan halaman jurnal).
- [7]. S. A. Abaya, “Association Rule Mining based on Apriori Algorithm in Minimizing Candidate Generation,” *International Journal of Scientific & Engineering Research*, 2012. (tidak ada vol, no dan halaman jurnal).
- [8]. J. Yabing, “Research of an Improved Apriori Algorithm in Data Mining Association Rules,” *International Journal of Computer and Communication Engineering*, 2013. (tidak ada vol, no dan halaman jurnal).
- [9]. J. Singh, et al, “Improving Efficiency of Apriori Algorithm Using Transaction Reduction,” *International Journal of Scientific and Research Publication*, 2013. (tidak ada vol, no dan halaman jurnal).