

ANALISIS KAIDAH ASOSIASI ANTAR ITEM DALAM TRANSAKSI PEMBELIAN MENGGUNAKAN DATA MINING DENGAN ALGORITMA APRIORI (STUDI KASUS: MINIMARKET GUN BANDUNGAN, JAWA TENGAH)

Adyawangkar Katon Prasidya¹⁾ dan Charitas Fibriani²⁾

^{1,2)} Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana
Jl. Diponegoro 52-60, Salatiga 50711, Indonesia
email: 682013045@student.uksw.edu¹⁾, charitas.fibriani@staff.uksw.edu²⁾

ABSTRAK

Data-data transaksi pembelian di minimarket yang selama ini hanya disimpan sebagai arsip dapat dimanfaatkan untuk menjawab masalah pengadaan stok barang, penentuan strategi promosi, dan penataan barang. Solusi pemecahan masalah-masalah tersebut dapat diperoleh menggunakan algoritma apriori, yang dapat digunakan untuk membantu menemukan kaidah asosiasi dalam pembelian item di minimarket. Informasi mengenai kaidah asosiasi dalam transaksi pembelian konsumen dapat dimanfaatkan untuk melakukan pengadaan stok barang yang lebih tepat guna dengan melakukan pengadaan stok barang yang berimbang pada item-item yang sering dibeli secara bersamaan, membuat strategi promosi yang lebih potensial untuk mendongkrak penjualan dengan mengacu pada kombinasi item yang sering dibeli secara bersamaan, dan menata barang di minimarket dengan berorientasi pada item-item yang sering dibeli secara bersamaan. Penelitian ini bertujuan menemukan kaidah asosiasi dalam pembelian item-item di minimarket untuk memecahkan masalah pengadaan stok barang, penentuan strategi promosi, dan penataan barang di minimarket.

Kata Kunci: *apriori algorithm, association rule mining, data mining*

ABSTRACT

Transactional data in minimarket which is frequently used only for archive files can be used to address the issue of procurement of the stock of items, promotional strategy determination, and the arrangement of items. The solution proposed to solve these problems is using an apriori algorithm, which can be used to help discovering association rules in the purchase pattern of items in the minimarket. Information about the association rules in the purchase pattern of the consumer can be utilized to appropriately procure the stock of items by balancing the procurement of the stock of items on those items frequently purchased together, creating more potential promotional strategy to boost sales by referring to a combination of items which are often purchased together, and arranging items on minimarket oriented to the items that are often purchased together. This study aims to discover association rules in the purchase of the items in the minimarket to solve the issue of procurement of the stock of items, promotional strategy determination, and the arrangement of items.

Keywords: *apriori algorithm, association rule mining, data mining*

I. PENDAHULUAN

Toko atau minimarket rata-rata telah mengadopsi Teknologi Informasi dalam proses bisnisnya. Contoh adopsi Teknologi Informasi di toko atau minimarket adalah penggunaan aplikasi *desktop* untuk mempermudah proses pembelian. 10 dari 10 sampel minimarket yang diobservasi di kota Salatiga, Jawa Tengah telah menggunakan aplikasi *desktop* untuk memproses transaksi pembelian [1]. Data transaksi pembelian di toko akan bertambah setiap hari, dan membutuhkan ruang penyimpanan data yang besar. Data-data transaksi pembelian sering tidak digunakan untuk kepentingan lebih lanjut, melainkan hanya dijadikan arsip internal milik toko. Data transaksi pembelian yang tersimpan memiliki informasi-informasi yang bisa digali dengan teknik *data mining*, misalnya informasi mengenai kaidah asosiasi dalam pembelian konsumen dan *cluster-cluster* konsumen yang ada.

Informasi mengenai kaidah asosiasi dalam pola pembelian konsumen dapat digunakan oleh toko untuk mengatasi masalah seperti pengadaan stok barang yang tidak dapat diprediksi, penentuan strategi promosi yang tepat guna, dan penataan barang di toko atau minimarket. Masalah pengadaan stok barang yang tidak dapat diprediksi, penentuan strategi promosi yang belum dapat mendongkrak penjualan, dan penataan barang di toko atau minimarket yang belum berorientasi pada kecenderungan pola pembelian konsumen adalah masalah yang dihadapi oleh minimarket Gun yang berlokasi di Bandungan, Jawa Tengah, sehingga peneliti memutuskan untuk mengambil judul penelitian “Analisis Kaidah asosiasi antar Item dalam Transaksi Pembelian menggunakan Data Mining dengan Algoritma Apriori (Studi Kasus: Minimarket Gun Bandungan, Jawa Tengah)”.

Data Mining adalah salah satu cara untuk mengolah atau mencari informasi-informasi dalam sekumpulan data [2]. Ada banyak metode dalam Data Mining, salah satunya adalah algoritma apriori, yaitu algoritma yang digunakan untuk menghasilkan kaidah asosiasi, dengan pola “jika-maka”, untuk mengetahui kecenderungan pola pembelian konsumen, dan kaidah asosiasi antar satu item pembelian dengan item lainnya. Informasi mengenai kaidah asosiasi dalam transaksi pembelian konsumen dapat dimanfaatkan untuk melakukan pengadaan stok barang yang lebih tepat guna dengan melakukan pengadaan stok barang yang berimbang pada item-item yang sering dibeli secara bersamaan, membuat strategi promosi yang lebih potensial untuk mendongkrak penjualan dengan mengacu pada kombinasi item yang sering dibeli secara bersamaan, dan menata barang di minimarket dengan berorientasi pada item-item yang sering dibeli secara bersamaan. Algoritma dalam Association Rule Mining lainnya yang dapat digunakan untuk memperoleh kaidah asosiasi dalam sekumpulan data adalah Predictive Apriori Algorithm dan algoritma Tertius. Algoritma Apriori dipilih untuk memecahkan masalah, karena algoritma Apriori menghasilkan Association Rules yang lebih baik daripada Predictive Apriori Algorithm dan algoritma Tertius [3].

Penelitian ini bertujuan menganalisis data transaksi pembelian di minimarket Gun Bandungan, Jawa Tengah untuk mencari keterkaitan pembelian antar item untuk memecahkan masalah pengadaan stok barang yang tidak dapat diprediksi, penentuan strategi promosi yang belum dapat mendongkrak penjualan, dan penataan barang di toko atau minimarket yang belum berorientasi pada kecenderungan pola pembelian konsumen.

II. KAJIAN PUSTAKA

Penelitian mengenai implementasi algoritma Apriori dalam analisis transaksi pembelian pernah dilakukan. Penelitian ini meneliti mengenai penggunaan algoritma Apriori, dimana salah satu implementasi algoritma Apriori yang paling populer adalah dalam analisis keranjang belanja (*market basket analysis*) dan *cross selling programs*. Penelitian ini menyimpulkan bahwa algoritma Apriori merupakan algoritma yang sangat berguna untuk menemukan korelasi antar item yang tersembunyi di dalam sebuah *database*. Algoritma Apriori dapat dikombinasikan dengan teknik lain, seperti Rule Induction Technique untuk meningkatkan akurasi dari hasil perhitungan, sehingga pola pembelian pelanggan di minimarket dapat lebih dipahami [4].

Algoritma Apriori sering digunakan dalam analisis transaksi pembelian, atau disebut juga *market basket analysis* untuk menemukan kaidah asosiasi antar item yang dibeli, namun algoritma *apriori* juga dapat digunakan untuk menganalisis penyebab terjadinya kecelakaan. Penelitian ini melakukan analisis terhadap penyebab terjadinya kecelakaan, dengan mengkombinasikan algoritma Apriori dan AHP. Penelitian ini menyimpulkan bahwa faktor-faktor utama penyebab terjadinya kecelakaan adalah faktor pengemudi (berpengalaman atau tidak berpengalaman), keadaan jalan, keadaan lingkungan sekitar, dan kondisi kendaraan [5].

Penelitian lain mengenai algoritma Apriori bertujuan untuk memprediksi resiko penyakit jantung melalui analisis *frequent itemsets*. Penelitian ini menyimpulkan bahwa data-data rekam medis sangat penting, karena dengan *mining* data-data rekam medis dapat membantu diagnosa penyakit dan membantu mengambil keputusan dalam melakukan tindakan medis. Algoritma Apriori dapat digunakan untuk memprediksi pasien-pasien yang beresiko terkena penyakit jantung [6].

Analisis menggunakan algoritma Apriori, Simple K-Means, dan *clustering* dapat digunakan untuk menemukan pola tindak kriminal, meramal, dan memetakan jaringan kriminal dan mengidentifikasi tersangka. Penelitian ini menghasilkan beberapa kaidah asosiasi, mengenai kecenderungan tindak kriminal di beberapa titik rawan, jenis kelamin pelaku tindak kriminal yang paling sering muncul, dan status pernikahan dari rata-rata pelaku tindak kriminal [7].

Algoritma Apriori mudah untuk dipahami dan mudah untuk diimplementasi. Akan tetapi, algoritma Apriori memiliki kekurangan-kekurangan, yaitu algoritma Apriori melakukan *scan* terhadap database berkali-kali, sehingga algoritma Apriori memerlukan alokasi memori yang besar dan waktu yang banyak untuk memproses sebuah database yang berukuran besar. Penelitian ini melakukan pengembangan dari algoritma Apriori, dengan berbasis *Boolean matrix* dan *Hadoop*. Dengan algoritma baru yang diajukan, database transaksi diganti dengan *Boolean matrix*, sehingga *non-frequent itemsets* dapat dieliminasi dari matriks. Selain itu, algoritma ini tidak perlu melakukan *scan* terhadap database, namun cukup dengan menggunakan *Boolean matrix* dengan operasi vector “DAN”. Algoritma ini dapat meningkatkan efisiensi dari algoritma Apriori tradisional [8].

Algoritma apriori dapat diterapkan juga di sektor medis. Penelitian menyimpulkan bahwa algoritma apriori menghasilkan Association Rules yang lebih baik daripada Predictive Apriori Algorithm dan algoritma Tertius. Penelitian ini menemukan bahwa algoritma Apriori berguna dalam menemukan pola-pola tersembunyi dalam *database*, seperti keefektifan perlakuan medis, prediksi terjadinya wabah penyakit dalam suatu daerah, dan keterkaitan antar suatu penyakit dengan faktor-faktor lain seperti gender, usia, dan pekerjaan. Informasi-informasi yang diperoleh dapat digunakan untuk meningkatkan pengambilan keputusan dalam bidang medis, deteksi awal dari persebaran penyakit, dan mencegah berbagai macam penyakit [3].

Penelitian ini bertujuan menemukan kaidah asosiasi dalam pembelian item-item di minimarket untuk melakukan

pengadaan stok barang yang lebih tepat guna dengan melakukan pengadaan stok barang yang berimbang pada item-item yang sering dibeli secara bersamaan, membuat strategi promosi yang lebih potensial untuk mendorong penjualan dengan mengacu pada kombinasi item yang sering dibeli secara bersamaan, dan menata barang di minimarket dengan berorientasi pada item-item yang sering dibeli secara bersamaan.

Sebagai contoh, ditemukan kaidah asosiasi pembelian konsumen antara item A dan B, maka pengadaan stok item “A” dan “B” harus dilakukan secara berimbang, karena kedua item sering dibeli secara bersamaan. Jika stok item “A” mulai menipis, stok item “B” juga harus ditambah jika memungkinkan, agar ketiadaan persediaan salah satu atau kedua item bisa dihindari. Strategi promosi yang akan dibuat untuk mendorong penjualan dapat diterapkan pada salah satu dari 2 item tersebut (A atau B), misalnya promosi produk “A” dapat ditawarkan kepada konsumen yang membeli produk “B”. Strategi promosi dengan mengadakan paket penjualan kedua produk “A” dan “B” tidak akan meningkatkan pembelian konsumen [9]. Strategi promosi yang baik akan mendorong pelanggan untuk membeli item lebih banyak dari biasanya. Penataan barang pada minimarket dapat dilakukan berdasarkan kaidah asosiasi yang ditemukan, yaitu dengan meletakkan item “A” di dekat “item “B”, sehingga pembelian salah satu item akan mendorong pembelian item yang lain. Item-item yang berkaitan harus diletakkan berdampingan atau dekat untuk mengingatkan pelanggan mengenai item lain yang memiliki kaitan dengan item yang dibeli sebelumnya, dan mendorong angka penjualan dari item-item yang berkaitan tersebut.

Data mining adalah sebuah proses pencarian korelasi, pola, dan tren, melibatkan bidang *machine learning*, statistik, dan teknik visualisasi melalui penyaringan sejumlah data yang besar. *Data mining* semakin populer dengan tren *big data* yang tidak terelakkan seiring aksesibilitas informasi yang semakin mudah dan ketersediaan informasi yang semakin meluas dari waktu ke waktu. *Data mining* juga semakin berkembang seiring adanya tren baru dalam dunia Teknologi Informasi untuk mengidentifikasi data-data berharga dan pengetahuan yang terkandung dalam suatu Sistem Informasi. Teridentifikasinya data dan pengetahuan dapat menjadi kunci untuk mengungguli kompetitor dalam dunia bisnis. Hal ini menimbulkan kebutuhan baru, yaitu kebutuhan akan adanya *tools* untuk menganalisis dan memodelkan data-data [10].

Sebelum menggunakan data untuk dianalisis dengan algoritma-algoritma *data mining*, ada beberapa tahapan yang harus dilakukan yaitu tahap *preprocessing data*. Tahap *preprocessing data* dilakukan karena tidak semua atribut dalam data akan digunakan untuk *mining data*, pengecekan data-data yang tidak lengkap, tidak konsisten, dan redundan. Tahapan *preprocessing data* dalam *data mining* adalah *data cleansing*, *data integration*, *data transformation*, dan *data reduction*. *Data cleansing* bertujuan untuk mengecek dan membenahi data-data yang tidak lengkap, tidak konsisten, dan redundan. Absennya tahap *data cleansing* dapat mengakibatkan hasil *mining data* yang didapatkan tidak akurat. *Data integration* bertujuan untuk menyatukan data-data yang memiliki atribut berbeda, namun arti sama, misalnya *customer_id* dan *cust_id*, yang dapat mengakibatkan inkonsistensi dan redundansi data. *Data transformation* bertujuan untuk melakukan normalisasi data yang belum ternormalisasi. Data yang belum ternormalisasi akan mengakibatkan hasil *mining data* tidak akurat. *Data reduction* bertujuan membuat ukuran data yang akan dianalisis menjadi lebih kecil, dengan beberapa metode seperti *data aggregation*, *attribute subset reduction*, *dimensionality reduction*, dan *numerosity reduction*. Walau data yang dianalisis berukuran lebih kecil, hasil analitik yang didapatkan dari proses *data mining* akan sama (atau hampir sama) [2].

Beberapa fungsi-fungsi umum *data mining* adalah *mining frequent pattern*, asosiasi, klasifikasi, prediksi, *outlier analysis*, dan *evolution analysis*. *Mining frequent pattern*, adalah proses pencarian pola-pola menarik, tak terduga, dan berguna dalam sekumpulan data. Asosiasi, adalah proses menemukan kaidah keterkaitan antara suatu kombinasi item dalam sekumpulan data. Klasifikasi, adalah proses mengelompokkan data ke dalam kelas-kelas berbeda, dimana kumpulan data di setiap kelas adalah data-data yang memiliki kategori sama. Prediksi, adalah proses identifikasi tren atau data-data numerik yang akan ada di masa yang akan datang, dengan menganalisis data yang tersedia. *Outlier analysis*, adalah proses menemukan data-data yang menyimpang dari data-data lain. *Evolution analysis*, adalah proses mendefinisikan kebiasaan atau tren dari data yang senantiasa berubah seiring waktu [2].

Frequent patterns, adalah pola-pola yang sering muncul di dalam sekumpulan data. *Frequent patterns* ada banyak jenisnya, mulai dari *itemsets*, *subsequences*, dan *substructures*. *Frequent itemset* biasanya merujuk pada sekumpulan item yang sering muncul bersama-sama dalam *database*, misalnya roti dan susu. Item yang sering muncul bersama-sama disebut memiliki kaidah asosiasi atau kaidah keterkaitan, karena dalam transaksi di dalam *database*, kehadiran salah satu item memiliki hubungan dengan item-item lainnya. *Association rule mining* adalah teknik *data mining* untuk menemukan kaidah asosiasi atau keterkaitan antar suatu kombinasi item dalam sekumpulan data. Suatu kaidah asosiasi dikatakan menarik, jika memenuhi nilai minimal dari 2 parameter, yaitu *support* dan *confidence*. *Support* adalah persentase kombinasi item muncul di dalam *database*. *Confidence* adalah persentase kuatnya hubungan antar item [2].

Kaidah asosiasi dapat dinyatakan dalam bentuk:

$$\text{Komputer} \Rightarrow \text{Software} [\text{support} = 1\%, \text{confidence} = 50\%]$$

Confidence sebesar 50% artinya jika seorang pelanggan membeli computer, ada kemungkinan sebesar 50% pelanggan tersebut akan membeli *software* juga, atau dengan kata lain, dari 50% transaksi di *database* yang memuat item komputer, juga memuat item *software*. Sedangkan *support* sebesar 1% artinya sebesar 1% dari keseluruhan transaksi yang diteliti menunjukkan bahwa komputer dan *software* dibeli secara bersamaan [2].

Tujuan dari Association Rule Mining adalah menemukan semua kaidah asosiasi yang memenuhi syarat minimum untuk *support* (*minimum support*) dan syarat minimum untuk *confidence* (*minimum confidence*). Kaidah asosiasi yang memenuhi *minimum support* dan *minimum confidence* dikatakan sebagai kaidah asosiasi yang menarik (*interesting rule*) [2].

Algoritma Apriori adalah teknik *data mining* yang sering digunakan untuk menemukan kaidah asosiasi di dalam *database*. Algoritma Apriori paling sering digunakan dalam analisis keranjang belanja, untuk mengetahui item-item yang sering dibeli secara bersamaan oleh konsumen. Algoritma Apriori diajukan pertama kali oleh R. Agrawal dan R. Srikant pada tahun 1994. Algoritma Apriori menggunakan pendekatan berulang yang disebut dengan *level-wise search*, dimana *n-itemset* digunakan untuk mencari $(n+1)$ -*itemset*. Pertama-tama, sejumlah 1-*itemset* dicari dengan menelusuri *database* untuk menentukan jumlah kemunculan masing-masing item, dan mengumpulkan item-item yang memenuhi *minimum support*. Nilai *support* sebuah item diperoleh dengan persamaan 1.

$$\text{Support}(A) = \frac{\Sigma \text{transaksi mengandung } A}{\text{Total transaksi}} \quad (1)$$

Nilai *minimum support* ditentukan oleh analis. Jika kaidah asosiasi yang dicari adalah kaidah asosiasi yang memiliki keterkaitan yang kuat, maka ditetapkan *minimum support* dengan persentase yang tinggi. Item-item yang memenuhi *minimum support* yang ditentukan dinyatakan sebagai L_1 . Selanjutnya, L_1 digunakan untuk mencari L_2 , yaitu sekumpulan 2-*itemset*, dengan cara melakukan proses kombinasi atau *join* dari *itemset-itemset* yang ada. Selanjutnya, L_2 digunakan untuk mencari L_3 , dan seterusnya, hingga tidak ada lagi *itemset* yang memenuhi *minimum support*. Proses pencarian L_n *itemset* memerlukan penelusuran penuh dari *database*. Untuk mencari nilai *support* dari n item, digunakan persamaan 2.

$$\text{Support}(A \cap B \dots \cap n) = \frac{\Sigma \text{transaksi mengandung } A \cap \Sigma \text{transaksi mengandung } B \dots \cap \Sigma \text{transaksi mengandung } n}{\text{Total transaksi}} \quad (2)$$

Setelah menemukan kumpulan *itemset* yang sering muncul dari *database*, dibentuk kaidah asosiasi yang memenuhi syarat *minimum confidence* yang ditetapkan, dengan menghitung *confidence* tiap *itemset*, dengan persamaan 3.

$$\text{Confidence}(A \Rightarrow B) = P(B|A) = \frac{\Sigma \text{transaksi mengandung } A \text{ dan } B}{\Sigma \text{transaksi mengandung } A} \quad (3)$$

Algoritma Apriori mudah dipahami dan diimplementasi, namun algoritma Apriori memerlukan waktu yang banyak, karena algoritma Apriori melakukan penelusuran *database* secara menyeluruh untuk menemukan tiap kombinasi *itemset*.

III. METODE PENELITIAN

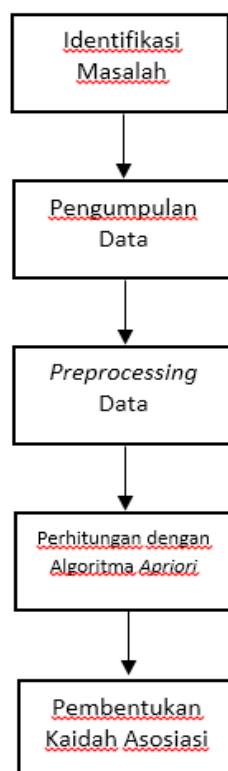
Analisis kaidah asosiasi antar item di minimarket Gun Bandungan, Jawa Tengah akan diselesaikan melalui beberapa tahapan-tahapan. Ada 5 tahapan yang dilakukan dalam penelitian, yaitu (1) Identifikasi Masalah, (2) Pengumpulan Data, (3) *Preprocessing Data*, (4) Perhitungan dengan Algoritma Apriori, dan (5) Pembentukan Kaidah Asosiasi. Data transaksi pembelian yang digunakan dalam penelitian adalah data transaksi pembelian yang terjadi dalam kurun waktu 6 hari operasional minimarket (10-Oktober-2016 s/d 15-Oktober-2016).

Tahap pertama dalam penelitian adalah melakukan identifikasi masalah yang terdapat pada minimarket Gun Bandungan, Jawa Tengah. Masalah-masalah yang terdapat pada minimarket Gun Bandungan, Jawa Tengah adalah pengadaan stok barang yang tidak dapat diprediksi, penentuan strategi promosi yang belum dapat mendorong penjualan, dan penataan barang di toko atau minimarket yang belum berorientasi pada kecenderungan pola pembelian konsumen. Stok barang di minimarket Gun tidak dapat diprediksi, beberapa produk barang sering mengalami defisit stok, namun terjadinya defisit stok tidak menentu dan tidak dapat diperkirakan. Promosi yang diadakan di minimarket Gun terbukti tidak dapat meningkatkan penjualan secara signifikan, karena dalam penentuan promosi minimarket tidak memiliki standar yang tepat atau terarah. Penataan barang di minimarket Gun dipisahkan berdasarkan kategori-kategori produk yang ada, namun minimarket tidak memanfaatkan penataan

barang berdasarkan kecenderungan pembelian konsumen dalam item-item yang sering dibeli secara bersamaan. Masalah-masalah yang ada akan dipecahkan menggunakan metode *data mining* dengan menggunakan algoritma Apriori. Metode yang digunakan untuk mengidentifikasi masalah adalah wawancara dengan pihak manajemen minimarket Gun Bandungan, Jawa Tengah.

Tahap kedua adalah pengumpulan data transaksi pembelian. Data transaksi pembelian akan digunakan untuk analisis kaidah asosiasi dalam pembelian barang oleh konsumen, dengan teknik *data mining* menggunakan algoritma Apriori. Metode yang digunakan untuk pengumpulan data adalah observasi, data transaksi diperoleh dari manajemen minimarket Gun. Data yang digunakan dalam penelitian adalah data kuantitatif, berupa data transaksi pembelian yang terjadi dalam kurun waktu 6 hari operasional minimarket (10-Oktober-2016 s/d 15-Oktober-2016). Data yang digunakan adalah data sekunder.

Tahap ketiga adalah melakukan *preprocessing data*. Data dalam dunia nyata harus melalui proses *preprocessing data* sebelum dilakukan proses *data mining*, karena data-data di dunia nyata cenderung tidak lengkap (ada *field-field* yang masih kosong), terdapat *noise* atau *error*, dan tidak konsisten. *Preprocessing data* dilakukan untuk menyeleksi atribut data yang akan digunakan dalam *data mining*, serta melakukan pengecekan data-data transaksi pembelian yang tidak lengkap, tidak konsisten, dan redundan. Tahapan-tahapan dalam *preprocessing data* adalah *data cleansing*, *data integration*, *data transformation*, dan *data reduction*. *Data cleansing* bertujuan memeriksa dan membenahi data-data transaksi pembelian data-data yang masih kosong, tidak konsisten, dan redundan. Tahap *data cleansing* dilakukan dengan bantuan perangkat lunak Microsoft Excel. Microsoft Excel memiliki fitur yang dapat digunakan untuk membantu menemukan data-data transaksi yang masih kosong dan redundan. Fitur *conditional formatting* dalam Microsoft Excel digunakan untuk menyorot data-data transaksi pembelian yang redundan, dengan menggunakan fungsi “*Format only unique or duplicate values*”. Data-data transaksi yang masih kosong dapat ditemukan dengan menggunakan fungsi dari fitur *conditional formatting* yang lain, yaitu “*Format only cells that contain blanks*”. *Data integration* berfungsi mengintegrasikan data-data yang belum terintegrasi, misalnya nama kolom yang berbeda namun memiliki arti sama. Data transaksi pembelian tidak perlu diintegrasikan lebih lanjut, karena data transaksi pembelian tidak memuat atribut-atribut yang perlu dipangkas karena tidak diperlukan dalam analisis. *Data transformation* bertujuan untuk melakukan normalisasi data pada data yang belum dinormalisasi. Di dunia nyata, data-data yang akan melalui proses *data mining* biasanya berasal dari banyak sumber data. Data-data yang berasal dari beberapa sumber data berbeda kemungkinan memiliki data yang tidak koheren jika digabungkan. Oleh sebab itu, perlu dilakukan *data transformation*. Data yang belum ternormalisasi akan mengakibatkan inakurasi hasil perhitungan dengan algoritma *apriori*. Data transaksi pembelian tidak perlu ditransformasi, karena data transaksi pembelian telah berada dalam kondisi ternormalisasi. *Data reduction*



Gambar 1. Tahapan kegiatan

bertujuan membuat ukuran data yang akan dianalisis menjadi lebih kecil jika memungkinkan, namun tidak mengubah hasil analisis dengan algoritma Apriori. *Data reduction* dapat dilakukan dengan menggunakan beberapa metode, seperti *data aggregation*, *attribute subset reduction*, *dimensionality reduction*, dan *numerosity reduction*. Data transaksi pembelian tidak melewati proses *data reduction*, karena tidak memungkinkan.

Tahap keempat adalah melakukan perhitungan dengan algoritma *apriori*. Tahapan dalam algoritma *apriori* adalah menentukan *minimum support* dan *minimum confidence* untuk menentukan seberapa kuat kaidah asosiasi yang akan dicari, mencari *frequent itemset* yang memenuhi *minimum support* yang telah ditetapkan, dan menghitung *confidence* dari masing-masing *frequent itemset*, lalu mencari *frequent itemset* yang memenuhi *minimum confidence* yang telah ditetapkan. Nilai *support* dari kaidah asosiasi $A \Rightarrow B$ adalah perbandingan transaksi dalam *database* yang memuat kedua item, A dan B. Nilai *support* dapat dihitung dengan persamaan 2.

Nilai *confidence* dari kaidah asosiasi $A \Rightarrow B$ adalah nilai akurasi dari kaidah asosiasi, yang ditentukan dari persentase transaksi dalam *database* yang mengandung A dan juga mengandung B. Nilai *confidence* dapat dihitung dengan persamaan 3.

Menetapkan nilai *minimum support* dan *minimum confidence* dilakukan sesuai kebutuhan dan data yang tersedia, misalnya untuk menemukan kaidah asosiasi dalam pembelian item di minimarket, ditetapkan *minimum support* sebesar 20% dan *minimum confidence* sebesar 70%, namun untuk mendeteksi kasus penipuan asuransi atau deteksi tindakan terorisme, ditetapkan *minimum support* sebesar 1% atau lebih kecil, karena data yang memuat indikasi penipuan asuransi atau tindakan terorisme berjumlah kecil dibandingkan keseluruhan data yang ada [11]. Pada saat menentukan *minimum support* dalam analisis data pembelian di *minimarket* Gun, awalnya ditetapkan nilai yang tinggi, misalnya 20%. Nilai *minimum support* akan dikurangi sedikit demi sedikit, hingga ditemukan *frequent 1-itemset* yang cukup untuk membentuk *frequent 2-itemset* atau lebih. *Itemset* adalah sekumpulan item, dan *n-itemset* adalah sebuah *itemset* yang memiliki n-buah *item*. Misalnya {A, B} adalah *2-itemset*, dan {C, D, E} adalah *3-itemset*. Frekuensi sebuah *itemset* adalah jumlah transaksi pembelian dalam *database* yang mengandung suatu *itemset* tertentu. *Frequent itemset* adalah *itemset* yang muncul dalam data transaksi pembelian sebanyak sejumlah minimal, dimana frekuensinya lebih dari nilai *minimum support* yang telah ditetapkan. Menentukan *1-itemset* atau kemunculan setiap item dalam *database* dapat dilakukan dengan menggunakan fitur dari perangkat lunak *Microsoft Excel*, yaitu *count of data* dengan terlebih dahulu memanfaatkan fungsi *pivot table*. Setelah menggunakan fitur *count of data*, dapat diidentifikasi jumlah kemunculan tiap item dalam *database* transaksi, sehingga nilai *support* masing-masing item dapat dihitung. Perhitungan nilai *support* masing-masing item dapat dilakukan dengan menggunakan fitur *formula* dari *Microsoft Excel*, dengan menggunakan persamaan 1. Setelah nilai *support* masing-masing item didapatkan, item-item yang memiliki nilai *support* lebih dari atau sama dengan nilai *minimum support* akan digunakan untuk membentuk kombinasi *2-itemset*, dan item-item yang memiliki nilai *support* dibawah nilai *minimum support* akan dieliminasi, karena item-item dengan nilai *support* dibawah nilai *minimum support* tidak akan digunakan lagi dalam perhitungan, karena data-data yang akan dicari adalah data-data pembelian yang sering muncul dalam transaksi. Kombinasi *2-itemset* didapatkan dengan cara melakukan kombinasi dari *1-itemset* yang memiliki nilai *support* lebih dari atau sama dengan nilai *minimum support*. Setelah mendapatkan kombinasi *2-itemset*, dilakukan lagi proses perhitungan masing-masing *support* pada tiap kombinasi *2-itemset*, begitu seterusnya hingga tidak ada lagi kombinasi *n-itemset* yang memenuhi nilai *minimum support*. Pada saat menentukan *minimum confidence* dalam analisis data pembelian di *minimarket* Gun, awalnya ditetapkan nilai yang tinggi, karena dalam menentukan kaidah asosiasi dalam pembelian item di *minimarket* kaidah asosiasi yang dicari adalah asosiasi yang kuat, karena kaidah asosiasi yang akan dicari adalah kaidah asosiasi dari item yang paling sering dibeli oleh pelanggan, dan kombinasi pembelian yang sering muncul dari kombinasi item yang paling sering dibeli oleh pelanggan. Pada awalnya ditetapkan nilai *minimum confidence* sebesar 60%. Nilai *minimum confidence* ditambah sedikit demi sedikit hingga ditemukan kaidah-kaidah asosiasi yang dinilai mampu merepresentasikan pola pembelian konsumen. Semakin rendah nilai *minimum support* dan *minimum confidence*, semakin banyak kaidah asosiasi yang dihasilkan, namun akurasi yang dihasilkan juga semakin rendah, karena kaidah asosiasi pembelian item yang lemah akan turut dimunculkan dalam hasil. Sebaliknya, semakin tinggi nilai *minimum support* dan *minimum confidence*, akurasi akan semakin tinggi, karena hanya item-item yang memiliki asosiasi kuat dalam pembelian konsumen yang akan dimunculkan dalam hasil perhitungan. Proses perhitungan *confidence* dapat dilakukan dengan menggunakan fitur *formula* dari *Microsoft Excel* dengan menggunakan persamaan 3.

Tahap kelima dalam penelitian adalah membentuk kaidah asosiasi dari hasil perhitungan yang telah dilakukan. Setelah semua *frequent itemsets* ditemukan, dapat langsung dibentuk kaidah asosiasi dari *frequent itemsets* yang ada. Kaidah asosiasi yang dicari adalah kaidah asosiasi yang memiliki asosiasi kuat, yaitu kaidah-kaidah asosiasi yang memenuhi nilai *minimum support* dan *minimum confidence*. *Frequent itemsets* adalah kombinasi *item-item* yang telah memenuhi nilai *minimum support*, sehingga dari *frequent itemsets* dapat langsung dicari *frequent itemsets* yang memenuhi nilai *minimum confidence*. Kaidah asosiasi yang telah didapatkan dapat digunakan sebagai

1	10/10/2016	CLEAR HA	MIE SEDA	CHARM UI	PANTENE	OLTO ULT	ABC SAUS	CHARM EX	KECAP AB	DJARUM 1	LIFEBUOY	CUSSON K	CLEAR ICE	RINSO MC	ATTACK M
2	10/10/2016	CLOSE UP	SUN LIGHT	TROPICAL	LIFEBUOY	LIFEBUOY	LUX PUTIH								
3	10/10/2016	LAURIER S	CUSSON KID	STRAWBERRY	45G										
4	10/10/2016	BIORE HEA	STARMILD	LAURIER S	SAMPOER	SUPER PEL	FORVITA S	KECAP ABC	MANIS REFFIL	100ML					
5	10/10/2016	MOLTO SC	WISMILAK	LUX VELVE	DANCOW	TROPICAL	BIMOLI 2 L	CLEAR HA	SUPERMI	LAURIER N	MIE SEDA	JHONSON	BIMOLI 1 L	LUX POWER	REJOICE C
6	10/10/2016	KECAP PIR	DOVE ANTI	LIFEBUOY	PANTENE	GULA PAS	KANSAS H	GG DJAJA	TROPICAL	CHARM N	BIORE DAI	BIMOLI 2 L	ABSOLUT		
7	10/10/2016	SUPERMI	GG FILTER	AROMA	MILO ACTI	SUN LIGHT	TROPICAL	SUN LIGHT	PALLMALL	KECAP SET	MIE SEDA	LIFEBUOY	ABC SAUS	BIORE REL	BIORE HEA
8	10/10/2016	DJARUM 7	BENTOEL S	MOLTO ULTRA	PINK 300ML										
9	10/10/2016	SUN LIGHT	DANCOW	PANTENE	SMOOTH SILKY										
10	10/10/2016	SARI MUR	NUVO FA	INDOMIE	GG KRETE	YURI HANI	GULA PAS	SUPER PEL	GREEN TE	DJARUM E	ARDATH	LA MERAH	LUCKY STR	BENTOEL S	BIORE REL
11	10/10/2016	ABC SAME	ABC SAUS	ATTACK S	FILMA 1 LT										
12	10/10/2016	SKM BENC	SOSRO	CELUP 30'S											
13	10/10/2016	DANCOW	BAGUS NII	CLOSE UP	SHINZUI	HANA									
14	10/10/2016	LUX MAGI	ATTACK S	MARLBOR	INDOFOO	MOLTO PE	SHINZUI K	REJOICE C	CLASMILD	AVOLUTION	MERAH				
15	10/10/2016	BIORE WH	GG FILTER	DANCOW	DOVE INTI	SIMAS PAI	GG PROFE	WISMILAK	AVOLUTIC	DJARUM 7	COUNTRY	BIORE DAI	LAURIER S	PROFIT MI	MOLTO SC
16	10/10/2016	GULA PAS	SUN LIGHT	BIORE HEA	SARI MUR	BIORE ACT	BAGUS NII	INDOFOO	LIFEBUOY	BUKRIM S	KISPRAY V	INDOMIE	SURF CLEA	SKM BENC	MILO 3 IN
17	10/10/2016	SUPER PEL	MOLTO PE	PEPSODEN	INDOMIE	FRAIS WEI	ABC SAME	BIMOLI 1 L	DJARUM 7	GREEN TE	AROMA	CLOSE UP	SKM INDOMILK	COKLAT 385G	
18	10/10/2016	CLEAR ICE	DJARUM 7	BIORE ACT	U MILD	H & S CLEA	CHARM EX	KECAP SET	SUN LIGHT	ABC SAUS	FILMA 2 LT				
19	10/10/2016	LUCKY STR	PANTENE	LA MENTH	SUPER PEL	DJARUM E	AVOLUTIC	SARIWANGI	CELUP 25'S						
20	10/10/2016	H & S CLEA	KAPAL AP	ABC SAUS	CHARM N	DJARUM N	KECAP SET	KECAP AB	U MILD						
21	10/10/2016	PEPSODEN	PEPSODEN	YURI HANI	BIORE EN	H & S SELE	DOVE ANTI	DANDRUF							
22	10/10/2016	KECAP PIR	BUKRIM S	SKM ULTR	AROMA										
23	10/10/2016	COUNTRY	INTERNASIONAL												
24	10/10/2016	SARI MUR	GULA PASIR	A 1 KG (PUTIH)											

Gambar 2. Sampel data transaksi pembelian

Row Labels	Count of DATA	Support
ABC SAUS TOMAT 340ML	40	12,232
KECAP SEDAAP MANIS 225ML	36	11,009
H & S CLEAN N BALANCE	28	8,5627
MILO ACTIGEN-E 300G	27	8,2569
INDOMIE GORENG	25	7,6453
GULA PASIR A 1 KG (PUTIH)	23	7,0336
PEPSODENT STAWBERRY BUBBLE 50G	22	6,7278
SARI MURNI 1 LT	22	6,7278
RINSO MOLTO 900G	21	6,422
DJARUM BLACK CAPPUCINO	21	6,422
WOW SEIUTA BUNGA 550G	20	6,1162
GG FILTER	19	5,8104
SIMAS PALMIA SACHET 200G	19	5,8104
INDOMIE AYAM BAWANG	19	5,8104
...		
ATTACK COLOUR 900G	8	2,4465
INDOFOOD SAMBAL EXTRA PEDAS 140ML	8	2,4465
LAURIER SUPER REGULER 20'S	8	2,4465
PALLMALL BIRU	8	2,4465
SUPERMI GOBANG	8	2,4465

Gambar 3. Jumlah pembelian tiap item dalam transaksi

acuan untuk melakukan pengadaan stok barang yang lebih tepat guna dengan melakukan pengadaan stok barang yang berimbang pada item-item yang sering dibeli secara bersamaan, membuat strategi promosi yang lebih potensial untuk mendongkrak penjualan dengan mengacu pada kombinasi item yang sering dibeli secara bersamaan, dan menata barang di minimarket dengan berorientasi pada item-item yang sering dibeli secara bersamaan.

Batasan masalah dalam penelitian ini adalah 1) Data transaksi yang digunakan adalah data transaksi yang terjadi dalam 6 hari operasional *minimarket* (10-Oktober-2016 s/d 15-Oktober-2016). 2) Item-item yang akan diteliti adalah item-item yang memenuhi nilai *minimum support* sebesar 6% dan *minimum confidence* sebesar 92%.

IV. HASIL DAN PEMBAHASAN

Dataset yang digunakan adalah data transaksi pembelian di *minimarket* Gun dalam jangka waktu 10-Oktober-2016 s/d 15-Oktober-2016 yang disajikan pada Gambar 2. *Dataset* memuat informasi tiap transaksi yang terjadi.

Kombinasi 2-Itemset	Σ Transaksi	Support
ABC SAUS TOMAT 340ML , KECAP SEDAAP MANIS 225ML	21	6,422018349
ABC SAUS TOMAT 340ML , H & S CLEAN N BALANCE	21	6,422018349
ABC SAUS TOMAT 340ML , MILO ACTIGEN-E 300G	2	0,611620795
ABC SAUS TOMAT 340ML , INDOMIE GORENG	2	0,611620795
ABC SAUS TOMAT 340ML , GULA PASIR A 1 KG (PUTIH)	5	1,529051988
ABC SAUS TOMAT 340ML , PEPSODENT STAWBERRY BUBBLE 50G	4	1,22324159
ABC SAUS TOMAT 340ML , SARI MURNI 1 LT	5	1,529051988
ABC SAUS TOMAT 340ML , RINSO MOLTO 900G	4	1,22324159
ABC SAUS TOMAT 340ML , DJARUM BLACK CAPPUCINO	3	0,917431193
ABC SAUS TOMAT 340ML , WOW SEJUTA BUNGA 550G	6	1,834862385
KECAP SEDAAP MANIS 225ML , H & S CLEAN N BALANCE	21	6,422018349
KECAP SEDAAP MANIS 225ML , MILO ACTIGEN-E 300G	4	1,22324159
KECAP SEDAAP MANIS 225ML , INDOMIE GORENG	4	1,22324159
KECAP SEDAAP MANIS 225ML , GULA PASIR A 1 KG (PUTIH)	3	0,917431193
KECAP SEDAAP MANIS 225ML , PEPSODENT STAWBERRY BUBBLE 50G	3	0,917431193
KECAP SEDAAP MANIS 225ML , SARI MURNI 1 LT	3	0,917431193
KECAP SEDAAP MANIS 225ML , RINSO MOLTO 900G	3	0,917431193
KECAP SEDAAP MANIS 225ML , DJARUM BLACK CAPPUCINO	1	0,305810398
KECAP SEDAAP MANIS 225ML , WOW SEJUTA BUNGA 550G	7	2,140672783
H & S CLEAN N BALANCE, MILO ACTIGEN-E 300G	1	0,305810398
H & S CLEAN N BALANCE, INDOMIE GORENG	1	0,305810398
H & S CLEAN N BALANCE, GULA PASIR A 1 KG (PUTIH)	3	0,917431193
H & S CLEAN N BALANCE, PEPSODENT STAWBERRY BUBBLE 50G	2	0,611620795
H & S CLEAN N BALANCE, SARI MURNI 1 LT	3	0,917431193
H & S CLEAN N BALANCE, RINSO MOLTO 900G	3	0,917431193
H & S CLEAN N BALANCE, DJARUM BLACK CAPPUCINO	0	0
H & S CLEAN N BALANCE, WOW SEJUTA BUNGA 550G	5	1,529051988

(a)

MILO ACTIGEN-E 300G , INDOMIE GORENG	23	7,033639144
MILO ACTIGEN-E 300G , GULA PASIR A 1 KG (PUTIH)	1	0,305810398
MILO ACTIGEN-E 300G , PEPSODENT STAWBERRY BUBBLE 50G	2	0,611620795
MILO ACTIGEN-E 300G , SARI MURNI 1 LT	1	0,305810398
MILO ACTIGEN-E 300G , RINSO MOLTO 900G	2	0,611620795
MILO ACTIGEN-E 300G , DJARUM BLACK CAPPUCINO	1	0,305810398
MILO ACTIGEN-E 300G , WOW SEJUTA BUNGA 550G	4	1,22324159
INDOMIE GORENG , GULA PASIR A 1 KG (PUTIH)	1	0,305810398
INDOMIE GORENG , PEPSODENT STAWBERRY BUBBLE 50G	0	0
INDOMIE GORENG , SARI MURNI 1 LT	1	0,305810398
INDOMIE GORENG , RINSO MOLTO 900G	2	0,611620795
INDOMIE GORENG , DJARUM BLACK CAPPUCINO	1	0,305810398
INDOMIE GORENG , WOW SEJUTA BUNGA 550G	4	1,22324159
GULA PASIR A 1 KG (PUTIH), PEPSODENT STAWBERRY BUBBLE 50G	0	0
GULA PASIR A 1 KG (PUTIH), SARI MURNI 1 LT	21	6,422018349
GULA PASIR A 1 KG (PUTIH), RINSO MOLTO 900G	1	0,305810398
GULA PASIR A 1 KG (PUTIH), DJARUM BLACK CAPPUCINO	3	0,917431193
GULA PASIR A 1 KG (PUTIH), WOW SEJUTA BUNGA 550G	2	0,611620795
PEPSODENT STAWBERRY BUBBLE 50G , SARI MURNI 1 LT	1	0,305810398
PEPSODENT STAWBERRY BUBBLE 50G , RINSO MOLTO 900G	0	0
PEPSODENT STAWBERRY BUBBLE 50G , DJARUM BLACK CAPPUCINO	4	1,22324159
PEPSODENT STAWBERRY BUBBLE 50G , WOW SEJUTA BUNGA 550G	0	0
SARI MURNI 1 LT , RINSO MOLTO 900G	1	0,305810398
SARI MURNI 1 LT , DJARUM BLACK CAPPUCINO	3	0,917431193
SARI MURNI 1 LT , WOW SEJUTA BUNGA 550G	2	0,611620795
RINSO MOLTO 900G , DJARUM BLACK CAPPUCINO	0	0
RINSO MOLTO 900G , WOW SEJUTA BUNGA 550G	1	0,305810398
DJARUM BLACK CAPPUCINO , WOW SEJUTA BUNGA 550G	1	0,305810398

(b)

Gambar 4. Kombinasi 2 itemset.

(a) bagian 1, (b) bagian 2

Kombinasi 3-itemset	Jumlah Transaksi	Support
ABC SAUS TOMAT 340ML, KECAP SEDAAP MANIS 225ML, H & S CLEAN N BALANCE	21	6,422018349

Gambar 5. Kombinasi 3 itemset

Tiap baris dalam *dataset* memuat informasi mengenai item-item yang dibeli dalam 1 kali transaksi oleh 1 pelanggan. Misalnya, dalam 1 baris transaksi memuat data {GG FILTER, INDOMIE AYAM BAWANG, INDOMIE GORENG}, artinya 3 item tersebut adalah *item-item* yang dibeli oleh pelanggan dalam 1 kali transaksi. *Dataset* memuat total 327 transaksi.

Tahapan pertama dalam perhitungan adalah mencari *1-itemset* dengan cara menelusuri *database* untuk menentukan jumlah kemunculan masing-masing item, dan mengumpulkan item-item yang memenuhi *minimum support*. Menghitung *minimum support* dari masing-masing item dilakukan dengan menggunakan persamaan 1.

Minimum support yang ditetapkan adalah sebesar 6%, dan transaksi pembelian yang diteliti berjumlah 327 transaksi, sehingga item-item yang dinyatakan sebagai *frequent 1-itemset* adalah item-item yang muncul dalam transaksi pembelian sebanyak $327 * 6\% = 19,62 \approx 20$ kali.

Berdasarkan 246 jenis item yang muncul dalam 327 transaksi pembelian dihitung masing-masing nilai *support* dari tiap item, untuk mencari *frequent 1-itemset* yang memenuhi nilai *minimum support* dan mengeliminasi item-item yang tidak memenuhi nilai *minimum support*. Item yang memenuhi *minimum support* sebesar 6%, atau memenuhi minimal 20 kali transaksi pembelian berjumlah 11 item. Berdasarkan Gambar 2, item yang ditetapkan *frequent 1-itemset* atau L_1 adalah:

$$L_1 = \{ \{ \text{ABC Saus Tomat 340ml} \}, \{ \text{Kecap Sedaap Manis 225ml} \}, \{ \text{H \& S Clean N Balance} \}, \{ \text{Milo Actigen-E 300g} \}, \{ \text{Indomie Goreng} \}, \{ \text{Gula Pasir A 1 Kg (Putih)} \}, \{ \text{Pepsodent Stawberry Bubble 50g} \}, \{ \text{Sari Murni 1 Lt} \}, \{ \text{Rinso Molto 900g} \}, \{ \text{Djarum Black Cappucino} \}, \{ \text{Wow Sejuta Bunga 550g} \} \};$$

Setelah menetapkan L_1 , dicari L_{n+1} , atau L_2 . *Frequent 2-itemset* atau L_2 didapat dari proses penggabungan dengan L_{n-1} , yaitu L_1 . Kombinasi yang terbentuk dari proses penggabungan L_1 adalah ditunjukkan pada Gambar 4(a) dan 4(b).

Proses penggabungan L_1 menghasilkan 55 kombinasi *2-itemset*. Berdasarkan 55 kombinasi *2-itemset*, dicari lagi *support* dari masing-masing kombinasi item menggunakan persamaan 1 dengan cara menelusuri *database* untuk mengeliminasi atau memangkas *frequent 2-itemset* yang tidak memenuhi *minimum support*. Hasil penelusuran *database* menyimpulkan bahwa terdapat 5 kombinasi *itemset* yang memenuhi *minimum support* sebesar 6% atau memiliki jumlah transaksi lebih dari atau sama dengan 20 transaksi. Di dalam seluruh kombinasi *2-itemset* yang berjumlah 55 kombinasi hanya terdapat 5 kombinasi *2-itemset* saja yang memenuhi nilai *minimum support*, sehingga dapat disimpulkan bahwa tidak semua kombinasi *itemset* dapat membentuk kaidah asosiasi, dan proses eliminasi penting dilakukan untuk memangkas kombinasi-kombinasi *item* yang tidak memiliki asosiasi kuat sesuai dengan *minimum support* dan *minimum confidence* yang diharapkan. Berdasarkan Gambar 2, *item* yang ditetapkan *frequent 2-itemset* atau L_2 adalah:

$$L_2 = \{ \{ \text{ABC Saus Tomat 340ml, Kecap Sedaap Manis 225ml} \}, \{ \text{ABC Saus Tomat 340ml, H \& S Clean N Balance} \}, \{ \text{Kecap Sedaap Manis 225ml, H \& S Clean N Balance} \}, \{ \text{Milo Actigen-E 300g, Indomie Goreng} \}, \{ \text{Gula Pasir A 1 Kg (Putih), Sari Murni 1 Lt} \} \};$$

Setelah menetapkan L_2 , dicari L_{2+1} , atau L_3 . *Frequent 3-itemset* atau L_3 didapat dari proses penggabungan dengan L_{n-1} , yaitu L_2 . *Itemset-itemset* yang dapat digabung adalah *itemset-itemset* yang memiliki kesamaan dalam 1 *item* pertama. Berdasarkan L_2 , yang dapat dibentuk menjadi calon *frequent 3-itemset* adalah {ABC SAUS TOMAT 340ML,KECAP SEDAAP MANIS 225ML} dan {ABC SAUS TOMAT 340ML, H & S CLEAN N BALANCE}, sehingga ditetapkan:

Berdasarkan 1 kombinasi 3 item, dicari *support* dari kombinasi item tersebut menggunakan persamaan 1 dengan cara menelusuri *database*. Jika dari kombinasi 3 item tidak ada kombinasi yang memenuhi *minimum support*, maka tidak akan dibentuk L_3 , dan kaidah asosiasi yang akan dibentuk hanya akan ditetapkan dari L_2 . Hasil penelusuran

database menyimpulkan seluruh kombinasi 3-itemset memenuhi *minimum support* sebesar 6% atau memiliki jumlah transaksi lebih dari atau sama dengan 20. Oleh karena itu, ditetapkan L_3 :

$$L_3 = \{ABC \text{ Saus Tomat } 340\text{ml, Kecap Sedaap Manis } 225\text{ml, H \& S Clean N Balance}\}$$

Berdasarkan L_3 , tidak dapat lagi dibentuk L_4 , sehingga perhitungan mencari *frequent itemset* berhenti. Setelah menemukan semua *frequent itemset* dari database, dibentuk kaidah asosiasi yang memenuhi *minimum support* dan *minimum confidence*, dimana *minimum support* yang ditetapkan sebesar 6%, dan *minimum confidence* yang ditetapkan sebesar 92%. Perhitungan *support* dan *confidence* dilakukan dengan menggunakan persamaan 2 dan 3.

Kaidah asosiasi yang dibentuk dinyatakan dalam jika x, maka y, dimana x adalah *antecedent* dan y adalah *consequent*. Untuk membentuk kaidah asosiasi dari *frequent itemset* yang ada, dilakukan proses kombinasi untuk setiap kemungkinan yang ada [2]. Misalnya, data memuat *frequent itemset* $l = \{I1, I2, I3\}$, *frequent itemset* l mengandung *subset* $\{I1, I2\}, \{I1, I5\}, \{I2, I5\}, \{I1\}, \{I2\}$, dan $\{I5\}$. Kemungkinan kaidah asosiasi yang dapat dibentuk dari l adalah:

$$\begin{aligned} I1 \wedge I2 &\Rightarrow I5 \\ I1 \wedge I5 &\Rightarrow I2 \\ I2 \wedge I5 &\Rightarrow I1 \\ I1 &\Rightarrow I2 \wedge I5 \\ I2 &\Rightarrow I1 \wedge I5 \\ I5 &\Rightarrow I1 \wedge I2 \end{aligned} \tag{4}$$

Berdasarkan himpunan $L_2 = \{\{ABC \text{ SAUS TOMAT } 340\text{ML, KECAP SEDAAP MANIS } 225\text{ML}\}, \{ABC \text{ SAUS TOMAT } 340\text{ML, H \& S CLEAN N BALANCE}\}, \{KECAP SEDAAP MANIS } 225\text{ML, H \& S CLEAN N BALANCE}\}, \{MILO \text{ ACTIGEN-E } 300\text{G, INDOMIE GORENG}\}, \{GULA \text{ PASIR A } 1 \text{ KG (PUTIH), SARI MURNI } 1 \text{ LT}\}\}$, dapat disusun kaidah asosiasi seperti pada Gambar 6.

Berdasarkan 10 kaidah asosiasi yang terbentuk dari L_2 , ada 2 kaidah asosiasi yang memenuhi nilai *minimum confidence* sebesar 92%. L_2 yang terdiri dari 10 kaidah asosiasi hanya memuat 2 kaidah asosiasi yang memenuhi nilai *minimum confidence* yang ditetapkan. Dapat disimpulkan bahwa semakin tinggi nilai *minimum confidence*, maka kaidah asosiasi yang dihasilkan akan semakin sedikit, namun kaidah asosiasi yang dihasilkan akan semakin baik, karena menjamin hanya akan memuat kaidah asosiasi yang memuat keterkaitan antar *item* yang kuat, sehingga jika ingin mencari kaidah asosiasi yang banyak namun mengabaikan kuatnya keterkaitan antar *item* yang dimuat, nilai *minimum confidence* dapat ditetapkan di nilai yang rendah.

Itemset	Supp (%)	Conf (%)
ABC SAUS TOMAT 340ML => KECAP SEDAAP MANIS 225ML	6,422018349	52,5
KECAP SEDAAP MANIS 225ML => ABC SAUS TOMAT 340ML	6,422018349	58,33333333
ABC SAUS TOMAT 340ML => H & S CLEAN N BALANCE	6,422018349	52,5
H & S CLEAN N BALANCE => ABC SAUS TOMAT 340ML	6,422018349	75
KECAP SEDAAP MANIS 225ML => H & S CLEAN N BALANCE	6,422018349	58,33333333
H & S CLEAN N BALANCE => KECAP SEDAAP MANIS 225ML	6,422018349	75
MILO ACTIGEN-E 300G => INDOMIE GORENG	7,033639144	85,18518519
INDOMIE GORENG => MILO ACTIGEN-E 300G	7,033639144	92
GULA PASIR A 1 KG (PUTIH) => SARI MURNI 1 LT	6,422018349	91,30434783
SARI MURNI 1 LT => GULA PASIR A 1 KG (PUTIH)	6,422018349	95,45454545

Gambar 6. Kaidah asosiasi dari 2-itemset

Itemset	Supp (%)	Conf (%)
ABC SAUS TOMAT 340ML => KECAP SEDAAP MANIS 225ML , H & S CLEAN N BALANCE	6,42201835	52,5
KECAP SEDAAP MANIS 225ML , H & S CLEAN N BALANCE => ABC SAUS TOMAT 340ML	6,42201835	100
KECAP SEDAAP MANIS 225ML => ABC SAUS TOMAT 340ML , H & S CLEAN N BALANCE	6,42201835	58,33333333
ABC SAUS TOMAT 340ML , H & S CLEAN N BALANCE => KECAP SEDAAP MANIS 225ML	6,42201835	100
H & S CLEAN N BALANCE => ABC SAUS TOMAT 340ML , KECAP SEDAAP MANIS 225ML	6,42201835	75
ABC SAUS TOMAT 340ML , KECAP SEDAAP MANIS 225ML => H & S CLEAN N BALANCE	6,42201835	100

Gambar 7. Kaidah asosiasi dari 3-itemset

Berdasarkan himpunan $L_3 = \{ABC\ SAUS\ TOMAT\ 340ML, KECAP\ SEDAAP\ MANIS\ 225ML, H\ \&\ S\ CLEAN\ N\ BALANCE\}$, dapat dibentuk kaidah asosiasi seperti yang tertera di Gambar 7. Berdasarkan 6 kaidah asosiasi yang terbentuk dari L_3 , ada 3 kaidah asosiasi yang memenuhi nilai *minimum confidence* sebesar 92%.

Perhitungan *support* dan *confidence* pada Gambar 6 dan 7 menghasilkan kaidah asosiasi sebagai berikut:

- Jika membeli “INDOMIE GORENG” maka akan membeli “MILO ACTIGEN-E 300G” dengan *support* 7,03% dan *confidence* 92%.
- Jika membeli “SARI MURNI 1 LT” maka akan membeli “GULA PASIR A 1 KG (PUTIH)” dengan *support* 6,42% dan *confidence* 95,45%.
- Jika membeli “KECAP SEDAAP MANIS 225ML, H & S CLEAN N BALANCE” maka akan membeli “ABC SAUS TOMAT 340ML” dengan *support* 6,42 % dan *confidence* 100%.
- Jika membeli “ABC SAUS TOMAT 340ML, H & S CLEAN N BALANCE” maka akan membeli “KECAP SEDAAP MANIS 225ML” dengan *support* 6,42% dan *confidence* 100%.
- Jika membeli “ABC SAUS TOMAT 340ML, KECAP SEDAAP MANIS 225ML” maka akan membeli “H & S CLEAN N BALANCE” dengan *support* 6,42% dan *confidence* 100%.

Berdasarkan kaidah asosiasi yang dihasilkan, dapat diterapkan beberapa strategi untuk mendongkrak penjualan. Misalnya kemungkinan konsumen yang membeli “INDOMIE GORENG”, maka juga akan membeli “MILO ACTIGEN-E 300G”, maka produk “INDOMIE GORENG” akan diletakkan di rak yang berdekatan, sehingga pembelian salah satu item akan mendorong pembelian item yang lain. Selain itu, dapat juga diterapkan strategi promosi untuk salah satu dari 2 item tersebut, bukan dengan mengadakan paket promosi yang menggabungkan item “INDOMIE GORENG” dan “MILO ACTIGEN E-300G”, karena hal ini tidak akan meningkatkan pembelian konsumen. Promosi produk “INDOMIE GORENG” dapat ditawarkan kepada konsumen yang membeli item “MILO ACTIGEN E-300G” (teknik *cross selling*) [9]. Pengadaan stok “INDOMIE GORENG” dan “MILO ACTIGEN-E 300G” harus dilakukan secara berimbang, karena kedua item tersebut sering dibeli secara bersamaan. Contohnya, jika stok “INDOMIE GORENG” mulai menipis, stok “INDOMIE GORENG” harus ditambah, dan stok “MILO ACTIGEN-E 300G” juga harus ditambah jika memungkinkan, untuk menghindari ketiadaan persediaan kedua barang tersebut di *minimarket*.

V. KESIMPULAN

Berdasarkan hasil analisis data transaksi pembelian pada kurun waktu 10-Oktober-2016 s/d 15-Oktober-2016, dapat disimpulkan bahwa nilai *support* dan *confidence* berbanding lurus dengan kuatnya asosiasi antar item. *Frequent itemset* yang memenuhi *minimum support* dan *minimum confidence* digunakan untuk membentuk kaidah asosiasi. Analisis data transaksi pembelian menghasilkan beberapa kaidah asosiasi dalam pola pembelian item oleh pelanggan di *minimarket* Gun Bandungan, Jawa Tengah. Kaidah asosiasi yang dihasilkan dapat digunakan *minimarket* untuk menyusun strategi promosi berdasarkan item-item yang sering dibeli secara bersamaan untuk mendongkrak penjualan dengan melakukan teknik *cross selling* sebagai ganti dari strategi promosi dengan mengadakan *bundling item*, melakukan pengadaan stok barang yang lebih tepat guna dengan melakukan pengadaan stok barang yang berimbang pada item-item yang sering dibeli secara bersamaan, dan mengatur tata letak item di *minimarket* sesuai pola pembelian pelanggan untuk menaikkan angka penjualan dengan cara menata item-item yang sering dibeli bersamaan berdekatan.

Perhitungan data transaksi pembelian menggunakan algoritma *apriori* memiliki kendala, yaitu kendala optimisasi algoritma dan optimisasi waktu, dimana algoritma *apriori* membutuhkan penelusuran *database* secara menyeluruh untuk mendapatkan tiap kombinasi *itemset*, terutama dalam menganalisis data yang berukuran besar. Untuk meningkatkan algoritma *apriori*, dapat digunakan beberapa teknik, seperti *hash-based*, *transaction reduction*, *partitioning*, *sampling*, dan *dynamic itemset counting* [2].

DAFTAR PUSTAKA

- [1] Hasil observasi 10 *minimarket* di kota Salatiga, Jawa Tengah, Indonesia.
- [2] J. Han, dan M. Kamber, “Mining Frequent Patterns, Associations, and Correlations” dalam *Data Mining: Concept and Techniques*, edisi ke-2, San Fransisco, California, 2006, bab 5, hal. 227-272.
- [3] J. Divya dan G. S. Maniata. (November 2013). Implementation of Apriori Algorithm in Health Care Sector: A Survey. *International Journal of Computer Science and Communication Engineering*. [Online]. 2(4), hal. 26-32. Tersedia: <http://static.ijcsce.org/wp-content/uploads/2013/12/IJCSCE110513.pdf>
- [4] M. Kaur, S. Kang. (Juni 2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. *Procedia Computer Science*. [Online]. 85(1), hal. 78-85. Tersedia: <http://www.sciencedirect.com/science/article/pii/S1877050916305208>

- [5] J. Xi, Z. Zhao, W. Li, Q. Wang. (Februari 2016). A Traffic Accident Causation Analysis Method Based on AHP Apriori. *Procedia Engineering*. [Online]. *137(1)*, hal 680-687. Tersedia: <http://www.sciencedirect.com/science/article/pii/S1877705816003325>
- [6] M. Ilayaraja, T. Meyyapan. (November 2015). Efficient Data Mining Method to Predict the Risk of Heart Diseases through Frequent Itemsets. *Procedia Computer Science*. [Online]. *70(1)*, hal. 586-592. Tersedia: <http://www.sciencedirect.com/science/article/pii/S1877050915032044>
- [7] Z. Zakaria Suliman dan M. Ayman Altaher. (Desember 2013). Crime Data Analysis Using Data Mining Techniques to Improve Crime Prevention. *International Journal of Computers*. [Online]. *8(1)*, hal. 39-45. Tersedia: https://www.researchgate.net/publication/259477161_Using_Data_Mining_Techniques_to_Analyze_Crime_Patterns_in_the_Libyan_National_Crime_Data
- [8] H. Yu, J. Wen, H. Wang, L. Jun. (Desember 2011). An Improved Apriori Algorithm Based On the Boolean Matrix and Hadoop. *Procedia Engineering*. [Online]. *15(1)*, hal 1827-1831. Tersedia: <http://www.sciencedirect.com/science/article/pii/S1877705811018418>
- [9] A. Ng, "Association Rules and the Apriori Algorithm: A Tutorial", tidak dipublikasikan. Tersedia: <http://www.kdnuggets.com/2016/04/association-rules-apriori-algorithm-tutorial.html>
- [10] W. A. Aldana, "Introduction" dalam *Data Mining Industry: Emerging Trends and New Opportunities*, Cambridge, Massachusetts, 2000, hal. 8.
- [11] D.T. Larose, "Association Rules" dalam *Discovering Knowledge in Data: An Introduction to Data Mining*, Hoboken, New Jersey, 2005, bab 10, hal. 184.