

## Paper

### Penerapan Data Mining Dengan Algoritma Apriori Untuk Mengidentifikasi Aturan Asosiasi Dalam Transaksi Penjualan

Author: Abdul Jabbar Lubis, Kalvin Chiuloto



## Penerapan Data Mining Dengan Algoritma Apriori Untuk Mengidentifikasi Aturan Asosiasi Dalam Transaksi Penjualan

Abdul Jabbar Lubis<sup>1\*</sup>, Kelvin Chiuloto<sup>2</sup>

<sup>1,2,3</sup>Universitas Harapan Medan, Medan, Indonesia

<sup>1</sup>[abduljabbarlubis@unhar.ac.id](mailto:abduljabbarlubis@unhar.ac.id),

### Abstrak

Data mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Setiap perusahaan atau bidang usaha bisnis pasti mempunyai data transaksi untuk setiap penjualan. Data transaksi tersebut setiap hari semakin banyak dan bertambah. Terkadang data transaksi tersebut hanya dibiarkan menumpuk tanpa ada tindakan yang lebih lanjut. Dengan bertambahnya jumlah data, maka untuk menganalisis data secara manual perlu digantikan dengan aplikasi yang berbasis komputer. Sehingga proses penganalisaan dapat dilakukan secara otomatis dan lebih mudah. Pemanfaatan data transaksi yang ada untuk menunjang kegiatan dalam pengambilan keputusan tidak cukup hanya dengan mengandalkan data operasional saja, diperlukan suatu analisis data untuk menggali potensi-potensi informasi yang ada. penelitian ini melakukan analisa data dengan menggunakan data mining dan algoritma A-Priori yang berfungsi untuk membentuk kandidat kombinasi item yang mungkin, lalu diuji apakah kombinasi tersebut memenuhi parameter support dan confidence minimum yang merupakan nilai ambang yang diberikan oleh user.

**Kata Kunci:** Aplikasi, Data Mining, Algoritma A-Priori, Aturan Asosiasi.

### Abstract

Data mining is a process of extracting data or filtering data by utilizing a large enough data set through a series of processes to obtain valuable information from the data. Every company or business must have transaction data for each sale. The transaction data is increasing every day. Sometimes the transaction data is just left to pile up without any further action. With the increasing amount of data, it is necessary to analyze the data manually with computer-based applications. So that the analysis process can be done automatically and more easily. Utilization of existing transaction data to support activities in decision making is not enough just to rely on operational data alone, a data analysis is needed to explore the potential of existing information. This research analyzes data using data mining and the A-Priori algorithm which functions to form a candidate combination of possible items, then tested whether the combination meets the minimum support and confidence parameters which are the threshold values given by the user.

**Keywords:** Application, Data Mining, A-Priori Algorithm, Association Rules.

## 1. PENDAHULUAN

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Istilah data mining memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Data mining, sering juga disebut sebagai Knowledge Discovery in Database [1]. Data mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Setiap perusahaan atau bidang usaha bisnis pasti mempunyai data transaksi untuk setiap penjualan. Data transaksi tersebut setiap hari semakin banyak dan bertambah. Terkadang data transaksi tersebut hanya dibiarkan menumpuk tanpa ada tindakan yang lebih lanjut. Dengan bertambahnya jumlah data, maka untuk menganalisis data secara manual perlu digantikan dengan aplikasi yang berbasis komputer. Sehingga proses penganalisaan dapat dilakukan secara otomatis dan lebih mudah.

Pemanfaatan data yang ada untuk menunjang kegiatan dalam pengambilan keputusan tidak cukup hanya dengan mengandalkan data operasional saja, diperlukan suatu analisis data untuk menggali potensi-potensi informasi yang ada. Para pengambil keputusan berusaha untuk memanfaatkan gudang data yang sudah dimiliki untuk menggali informasi yang berguna membantu dalam mengambil keputusan. Hal ini mendorong munculnya

cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola yang penting atau menarik dari data dalam jumlah besar yang disebut dengan data mining. Penggunaan teknik data mining diharapkan dapat memberikan pengetahuan-pengetahuan yang sebelumnya tersembunyi didalam gudang data sehingga menjadi informasi yang berharga.

Analisis asosiasi atau association rule mining adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Algoritma aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian data mining, untuk menghasilkan pengetahuan. Pengetahuan untuk mengetahui item-item belanja yang sering dibeli secara bersamaan dalam suatu waktu. Aturan asosiasi yang berbentuk “if...then...” atau “jika...maka...” merupakan pengetahuan yang dihasilkan dari fungsi aturan asosiasi [2]. Pencarian asosiasi berasal dari pengolahan data transaksi pembelian barang dari setiap pembeli, kemudian dicari hubungan antar barang-barang yang dibeli. Pencarian informasi ini hampir sama dengan mencari peluang kemunculan barang yang dibeli sesuai dengan kebiasaan berbelanja masyarakat dan jumlah transaksi yang ada. Proses pencarian asosiasi ini menggunakan algoritma apriori, yang berfungsi untuk membentuk kandidat kombinasi item yang mungkin, lalu diuji apakah kombinasi tersebut memenuhi parameter support dan confidence minimum yang merupakan nilai ambang diberikan oleh user.

Data mining atau sering disebut sebagai knowledge discovery in database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar. Keluaran data mining ini bisa dipakai untuk membantu pengambilan keputusan di masa depan [3].

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Istilah data mining memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Data mining, sering juga disebut sebagai Knowledge Discovery in Database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [4]. Dalam data mining ada beberapa teknik menganalisis data, salah satunya adalah dengan algoritma apriori. Algoritma apriori adalah algoritma untuk menemukan frekuensi dan kombinasi frekuensi itemsets yang dijalankan pada sekumpulan data, sedangkan teknik untuk menemukan hubungan antara satu item dengan item yang lain, dan antara pola kombinasi itemset yang satu dengan pola kombinasi itemsets yang lain disebut association rules. Algoritma Apriori adalah algoritma yang digunakan untuk menghasilkan aturan asosiasi dengan pola “if...then” [1].

Algoritma Apriori menggunakan pendekatan iteratif yang dibuat dengan level-wise search, dimana k-itemset produk digunakan untuk mengeksplorasi  $(k+1)$ -itemset produk atau  $(k-1)$ -itemset. Notasi yang digunakan dalam algoritma apriori antara lain [5]:

1.  $C_k$  adalah kandidat k-itemset, dimana k menunjukkan jumlah pasangan item.
2.  $L_k$  adalah large k-itemset.
3. D adalah database transaksi dimana  $|D|$  adalah banyaknya transaksi.

Tahapan yang dilakukan algoritma apriori untuk membangkitkan large item set adalah sebagai berikut :

1. Menelusuri seluruh record di database dan menghitung support count dari setiap item. Ini adalah kandidat 1-itemset,  $C_1$ .
2. Large 1-itemset  $L_1$  dibangun dengan menyaring  $C_1$  dengan support count yang lebih besar atau sama dengan minimum support untuk dimasukkan kedalam  $L_1$ .
3. Untuk membangun  $L_2$  algoritma apriori menggunakan proses join untuk menghasilkan  $C_2$ .
4. Dari  $C_2$ , 2-itemset yang memiliki support count yang lebih besar atau sama dengan minimum support akan disimpan ke  $L_2$ .
5. Proses ini diulang sampai tidak ada lagi kemungkinan kitemset. Contoh tahapan pembangkitan  $C_1, L_1, C_2, L_2, C_3, L_3$

## 2. METODE PENELITIAN

### 2.1 Analisis Sistem

Data mining adalah proses analisis untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database. Salah satu metode yang digunakan dalam data mining yaitu Association Rule dengan metode Apriori, dimana aturan asosiasi ingin memberikan informasi dalam bentuk hubungan if-then atau jika-maka dari data transaksi pembelian konsumen. Kinimiko Coffe merupakan salah satu usaha yang bergerak dibidang penjualan produk makanan dan minuman siap saji. Dalam melakukan proses penjualan Kinimiko Coffe memiliki arsip penjualan yang berupa nota penjualan berisi beberapa data item penjualan dalam sekali transaksi. Dalam penulisan tugas akhir ini akan dicari nilai support dan confidence dari data transaksi pembelian konsumen pada Kinimiko Coffe. Data transaksi nantinya akan di mining dengan menerapkan association rule sehingga didapatkan aturan asosiasi baru yang berfungsi untuk memberikan informasi yang berguna untuk mengetahui jenis minuman dan makanan apa saja yang biasanya dibeli oleh konsumen, sehingga dapat dijadikan patokan untuk menentukan persediaan atau stok barang, sehingga stok untuk item (makanan dan minuman) tersebut harus lebih banyak daripada item yang lain.

### 2.2 Preprocessing Data

Preprocessing data dimaksudkan untuk mempersiapkan data sebelum diproses dengan teknik data mining menggunakan algoritma Apriori. Tahap preprocessing data meliputi sumber data, data cleaning dan pemodelan data. Analisa data merupakan proses untuk menyiapkan data mentah dengan tujuan agar data tersebut siap untuk diolah selanjutnya dengan data mining. Pemahaman ini tentang pengumpulan data awal. Data yang akan digunakan untuk memenuhi tujuan bisnis dan tujuan data mining yang telah ditetapkan adalah data transaksi pembelian konsumen selama periode waktu tertentu pada Kinimiko Coffee.

Karena tidak semua data digunakan maka perlu dilakukan pembersihan data agar data yang akan diolah benar-benar relevan dengan yang dibutuhkan. Pembersihan ini penting guna meningkatkan performa dalam proses mining. Cara pembersihan dengan menghapus atribut yang tidak terpakai dan menghapus data-data yang tidak lengkap isiannya. Terdapat empat atribut pada data transaksi yang digunakan, yaitu :

1. Banyaknya  
Atribut “Banyaknya” merupakan atribut untuk menentukan jumlah yang dibeli dalam satu jenis menu yang sama.
2. Nama Barang  
Atribut “Nama Barang” merupakan atribut jenis-jenis menu apa saja yang dibeli konsumen.
3. Harga  
Atribut “Harga” merupakan atribut harga per jenis-jenis menu.
4. Jumlah  
Atribut “Jumlah” merupakan atribut perkalian dari atribut harga dan atribut banyaknya.

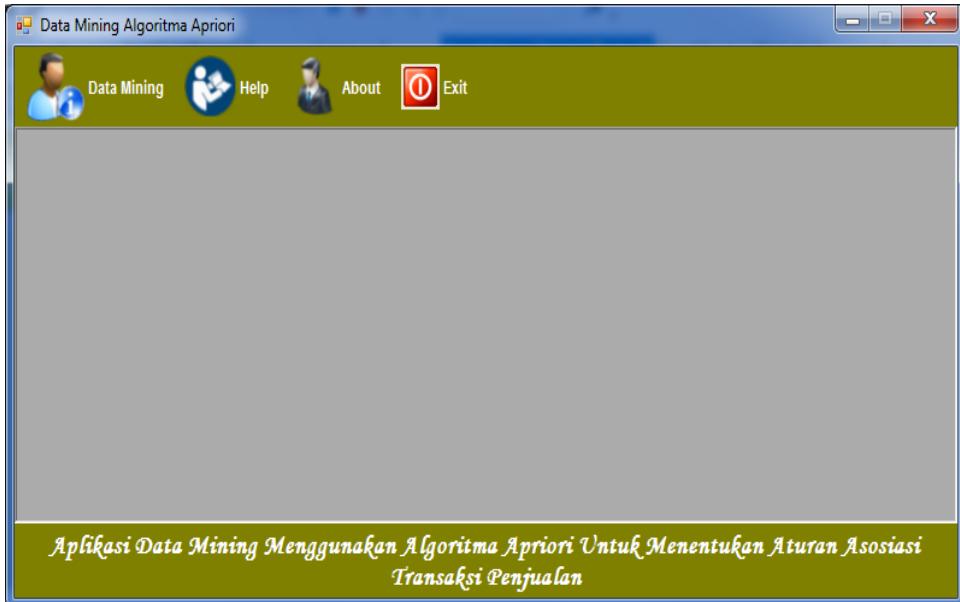
Data yang didapat dari Kinimiko Coffee berbentuk nota data kemudian dibersihkan dengan cara menghapus atribut-atribut yang tidak diperlukan untuk penelitian. Dalam hal ini hanya atribut Nama Barang yang akan dipakai dalam penelitian.

Teknik data mining yang digunakan dalam penelitian ini adalah teknik asosiasi menggunakan algoritma Association Rule. Pemodelan bertujuan untuk mencari aturan asosiasi yang selanjutnya dijadikan acuan untuk mengetahui makanan dan minuman apa saja yang sering dibeli oleh konsumen secara bersamaan. Langkah-langkah kerja dari algoritma Association Rule adalah sebagai berikut:

1. Tentukan besaran nilai minimum support dan confidence, dimana dalam penelitian ini jumlah besaran nilai minimum support = 8%, sedangkan besaran minimum confidence = 50%.
2. Scan database transaksi pembelian untuk mendapatkan 1-item set, kemudian ambil kandidat yang frekuensinya memenuhi minimum support count. Sedangkan yang tidak memenuhi minimum support dihilangkan atau dihapus.
3. Bangkitkan k+1 itemset dari itemset yang ada sebelumnya, kemudian buang atau bersihkan data kandidat yang subsetnya tidak memenuhi support count.
4. Hitung frekuensinya, ambil kandidat yang frekuensinya memenuhi dan selesai saat tidak ada kandidat yang memenuhi syarat.
5. Aturan asosiasi dihasilkan dengan menyeleksi yang memenuhi besaran minimum support dan minimum confidence.

### 3. HASIL DAN PEMBAHASAN

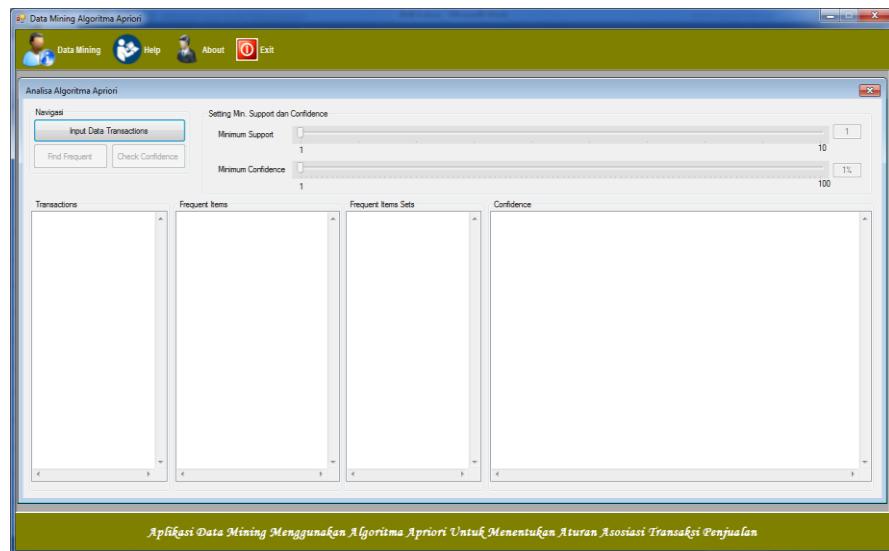
Implementasi sistem dilakukan dengan membuat file project dari perangkat lunak yang dibuat. File project merupakan suatu gabungan dari beberapa form program. Form berfungsi sebagai media tatap muka antara user dan sistem. Di samping itu juga, form juga berfungsi mempermudah user dalam menjalankan sistem. Setiap form program berisikan data tentang layar antar muka yang disimpan pada file form. Pada tahap ini, penulis melakukan pengujian terhadap kemampuan sistem untuk menggunakan form menu utama di mana user akan berinteraksi langsung dengan sistem ini. Adapun tampilan dari implementasi form Utama dapat dilihat pada gambar berikut.



Gambar 1. Implementasi Form Menu Utama

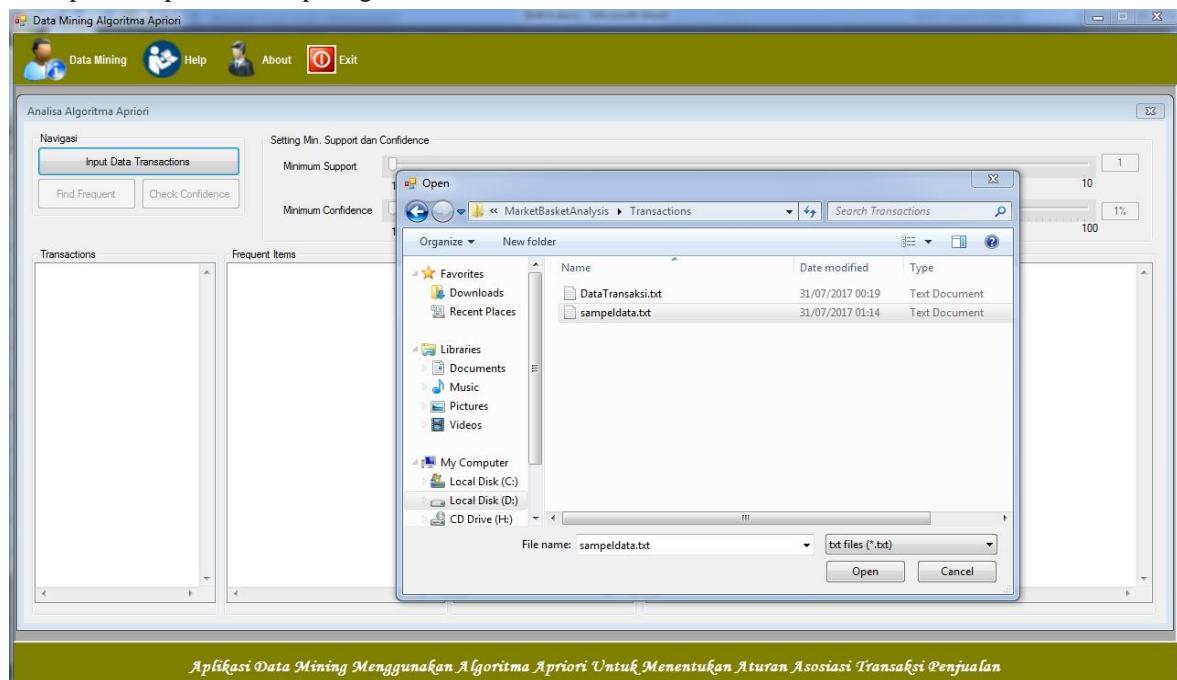
Form Menu Utama ini user melakukan interaksi yang mencakup empat buah menu, yaitu menu Data Mining yang digunakan untuk melakukan analisa data mining dengan algoritma A-Priori guna memperoleh pengetahuan tentang aturan asosiasi yang dihasilkan dari data transaksi. Menu Help yang digunakan untuk menampilkan informasi tata cara penggunaan aplikasi (manual book) kepada user, menu About yang digunakan untuk menampilkan informasi tentang pembuat program serta Menu Close yang berfungsi untuk keluar dari sistem.

Form Data Mining adalah sebuah form yang difungsikan untuk melakukan analisa data mining dengan algoritma A-Priori. Adapun implementasi tampilan awal dari Form Data Mining seperti terlihat pada gambar tersebut.



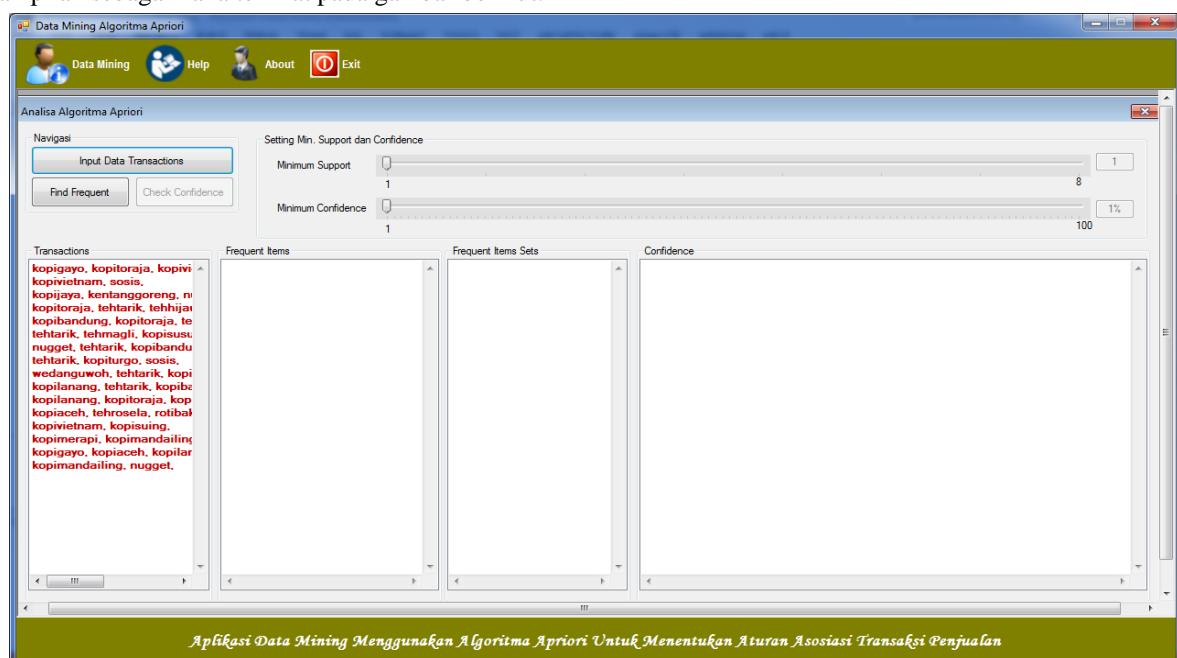
Gambar 2. Implementasi Awal Form Data Mining

Berdasarkan gambar 2 dapat dijelaskan Form Data Mining merupakan form yang difungsikan untuk melakukan analisa data mining dengan menerapkan algoritma A-Priori. Proses analisa pada form data mining dimulai dengan menginputkan data transaksi, yaitu user memilih tombol ‘Input Data Transaction’ dan sistem akan menampilkan seperti terlihat pada gambar berikut.



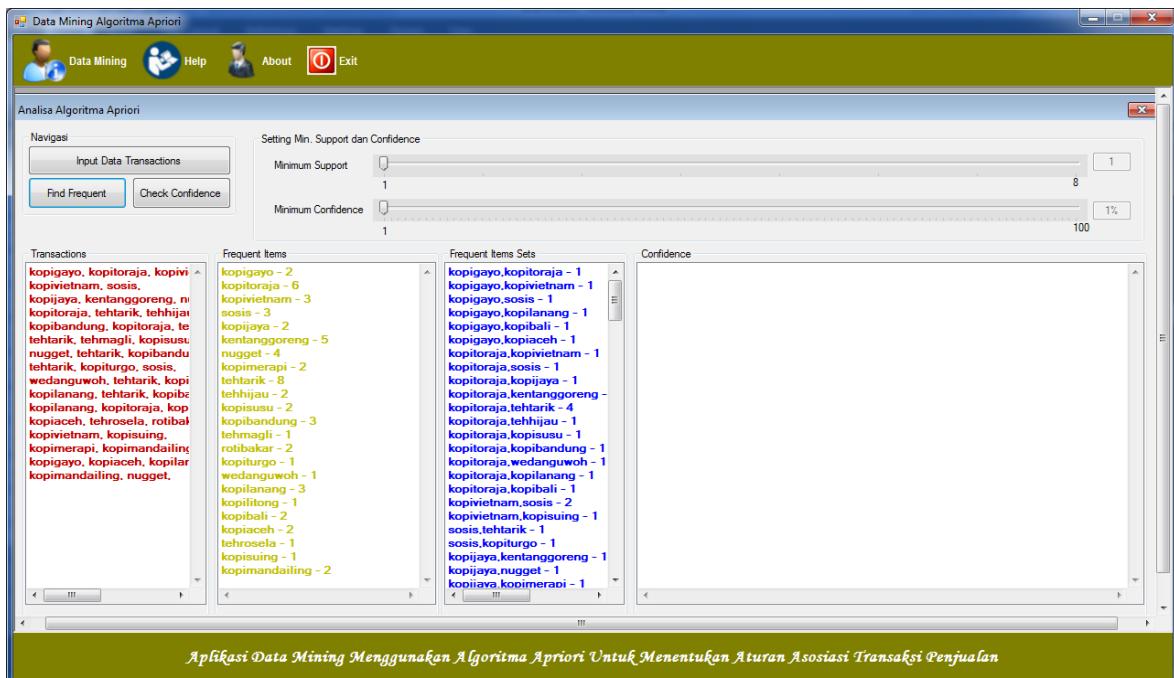
Gambar 3. Implementasi Pilih Data Transaksi

Berdasarkan gambar 4 dapat dijelaskan proses penginputan data transaksi dengan mengimport file .txt yang sudah disiapkan sebelumnya. Setelah data diimport maka sistem akan menampilkan ke form data mining dengan tampilan sebagaimana terlihat pada gambar berikut.



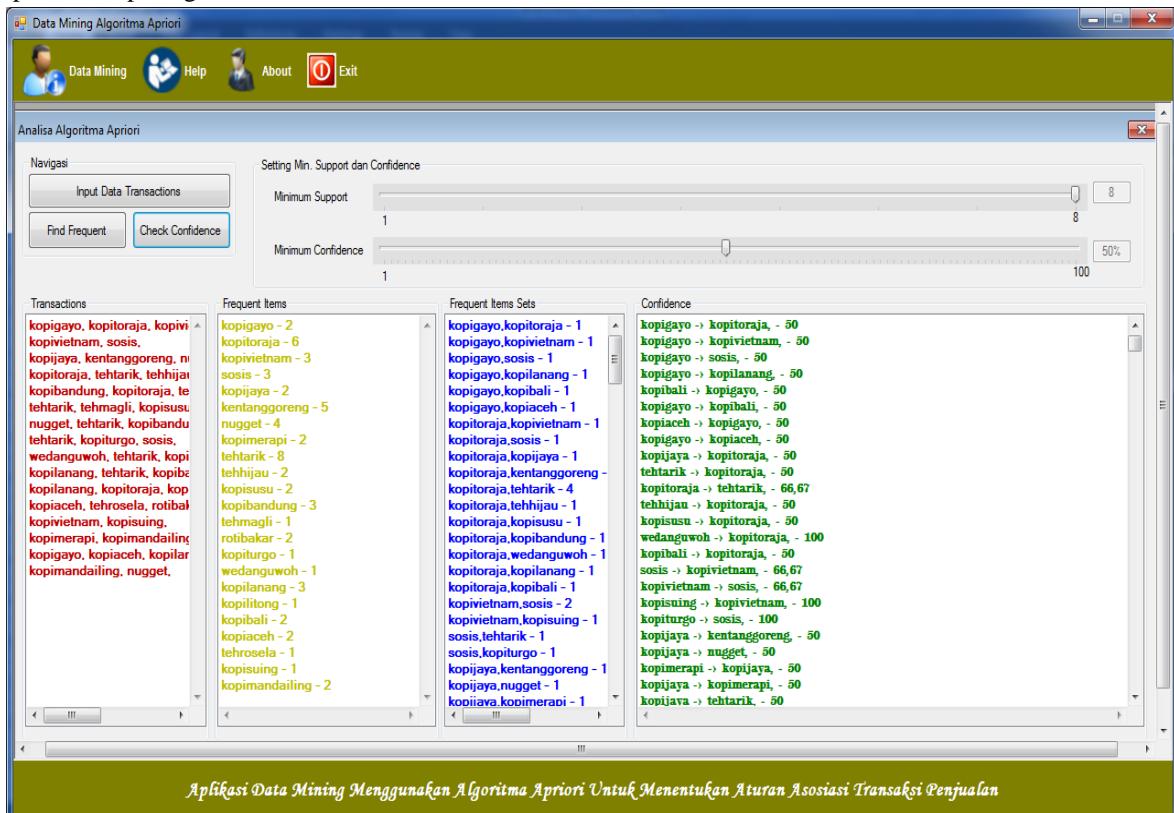
Gambar 4. Implementasi Tampil Data Transaksi

Berdasarkan gambar 4 data transaksi berhasil diimport ke dalam sistem. Langkah selanjutnya untuk melakukan analisa data mining yaitu user memilih tombol ‘Find Frequent’ maka sistem akan menampilkan form seperti terlihat pada gambar berikut.



Gambar 5. Implementasi Tampil Frequent Item Transactions

Berdasarkan gambar 5 dapat dijelaskan pada data transaksi sistem akan menampilkan frekuensi tem dan frekuensi item sets dari semua data transaksi. Tahap selanjutnya yaitu user mementukan nilai minimum support dan minimum confidence untuk menampilkan aturan asosiasi. Pada tahap implementasi in Penulis menentukan nilai minimum support sebesar 8% dan minimum confidence sebesar 50%. Adapun hasil aturan asosiasi yang dihasilkan dapat dilihat pada gambar berikut.



Gambar 6. Implementasi Hasil Association Rule

Berdasarkan gambar 6 dapat dijelaskan aturan asosiasi yang dihasilkan dari 16 data transaksi dengan minimum support sebesar 8% dan minimum confidence sebesar 50%, Maka dihasilkan 76 aturan asosiasi. Support disini artinya suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi. Sedangkan Confidence adalah suatu ukuran yang menunjukkan hubungan antar dua item secara conditional. Diantara aturan asosiasi yang dihasilkan yaitu ‘Jika membeli Kopi Gayo maka membeli Kopi Toraja dengan nilai confidence 50%’. Aturan tersebut berarti “seorang konsumen yang membeli Kopi Gayo punya kemungkinan 50% untuk juga membeli Kopi Toraja”. Aturan ini cukup signifikan karena mewakili 50% dari transaksi yang ada.

## 4. KESIMPULAN

Adapun kesimpulan dari hasil penelitian ini yaitu:

1. Teknik data mining dengan algoritma Apriori dapat diimplementasikan pada sistem transaksi penjualan untuk menghasilkan aturan asosiasi (association rule).
2. Penerapan data mining transaksi penjualan merupakan sebuah sistem pencarian aturan asosiasi melalui pengolahan data transaksi penjualan dari setiap konsumen yang bertujuan untuk mencari hubungan antara item yang sering dibeli bersamaan oleh konsumen dengan menghitung besaran nilai support dan confidence. Support disini artinya suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan data transaksi. Sedangkan confidence adalah suatu ukuran yang menunjukkan hubungan antar dua item secara conditional.
3. Dari hasil data mining transaksi penjualan didapatkan informasi item yang sering dibeli bersamaan dapat dijadikan patokan untuk menentukan persediaan atau stok barang, sehingga stok untuk item tersebut harus lebih banyak daripada item yang lain.
4. Dari hasil mining data transaksi penjualan dengan batasan minimum support sebesar 8% transaksi dan minimum confidence sebesar 50% membentuk 76 aturan asosiasi. Salah satu aturan asosiasi yang terbentuk yaitu ‘Jika membeli Kopi Gayo maka membeli Kopi Toraja dengan nilai confidence 50%’. Aturan tersebut berarti “seorang konsumen yang membeli Kopi Gayo punya kemungkinan 50% untuk juga membeli Kopi Toraja”. Aturan ini cukup signifikan karena mewakili 50% dari transaksi yang ada.

## DAFTAR PUSTAKA

- [1] M. . Aminullah, “Penerapan Data Mining Pada Penyewaan Film Di Atika Menggunakan Metode Association Rules,” Universitas Nusantara PGRI, 2016.
- [2] Budiono, “Penerapan Metode Association Rule Menggunakan Algoritma Apriori Untuk Mengidentifikasi Pola Penyakit Radang Sendi,” *Techno.COM*, vol. 13, no. 2, 2013.
- [3] Fadlina, “Data Mining Untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori (Studi Kasus Di Polsekta Medan Sunggal),” *Maj. Ilm. Inf. dan Teknol. Ilm.*, vol. 3, no. 1, 2014.
- [4] H. Marcos and I. Hidayah, “Implementasi Data Mining Untuk Klasifikasi Nasabah Kredit Bank”X“ Menggunakan Classification Rule,” in *Seminar Nasional Teknologi Informasi dan Multimedia*, Yogyakarta: STMIK AMIKOM Yogyakarta, 2014.
- [5] D. Astika and Nurdin, “Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Menggunakan Metode Apriori Pada Supermarket Sejahtera Lhokseumawe,” *Techsi*, vol. 6, no. 1, 2015.