

Implementasi Data Mining Pada Penjualan Toko Sembako Dengan Algoritma Apriori

1st Izz Abdur Rahman
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia
izzetabdur@student.telkomuniversity.ac.id

2nd Riska Yanu Farifah
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia
riskayanu@telkomuniversity.ac.id

3rd Faqih Hamami
Fakultas Rekayasa Industri
Universitas Telkom
Bandung, Indonesia
faqihhamami@telkomuniversity.ac.id

Abstrak

Penjualan toko sembako meningkat setiap bulannya. Semakin banyaknya jumlah toko sembako membuat para pengelola ingin membuat strategi pemasaran yang lebih baik. Untuk mengetahui penjualan terbanyak diperlukan algoritma apriori dengan bantuan tools Rapid Miner. Algoritma apriori ini akan membentuk frequent itemset dengan banyak yang telah ditentukan sebelumnya dengan dua parameter yaitu support dan confidence. Salah satu tahap analisis asosiasi algoritma yang efisien adalah analisis pola frekuensi tinggi. Support adalah presentase kombinasi item dalam database, sedangkan confidence adalah kuatnya hubungan antar barang dalam asosiasi. Algoritma apriori ini dapat membantu dalam data mining dan pemasaran. Penelitian menggunakan salah satu aplikasi data mining yaitu Rapid Miner. Penelitian ini bertujuan untuk memberikan gambaran tentang keterkaitan barang yang terjual di toko sembako dengan melihat data barang yang terjual selama 1 bulan. Adapun hasil yang didapatkan dari penelitian adalah nilai support 0.25% dan nilai confidence 0.5% dengan total dari keseluruhan data adalah 2328 data. Dengan adanya penelitian terhadap perhitungan pola penjualan barang di toko sembako tersebut, diharapkan pemilik toko sembako dapat melihat strategi bisnis yang lebih menguntungkan.

Kata Kunci : algoritma apriori, data mining, toko sembako, penjualan

Abstract

The sales of the grocery store increase every month. The increasing number of grocery stores makes managers want to make better marketing strategies. To find out the most sales, an apriori algorithm is needed with the help of the Rapid Miner tools. This apriori algorithm will form a frequent itemset with a predetermined number of parameters with two parameters, namely support and confidence. One of the stages of efficient algorithm association analysis is high-

frequency pattern analysis. Support is the percentage of item combinations in the database, while confidence is the strength of the relationship between items in the association. This apriori algorithm can help in data mining and marketing. The research uses one of the data mining application, namely Rapid Miner. This study aims to provide an overview of the interrelationships of goods sold in grocery store by looking at the data of good sold for 1 month. The results obtained from the research are the support value of 0.25% and the confidence value of 0.5% with total of 2328 data. With the research on the calculation of the pattern of sales of goods at the grocery store, it is hopes that the owner of the grocery store can see a more profitable business strategy.

Keywords: apriori algorithm, data mining, grocery store, sales.

I. PENDAHULUAN

Semakin banyaknya persaingan antara industri penjualan, menuntut para pengelola untuk menemukan suatu strategi yang bisa meningkatkan penjualan dan pemasaran. Untuk itu maka para pengelola harus mencermati pola pola pembelian yang dilakukan oleh konsumen. Dengan adanya penjualan setiap saat, semakin banyak data yang bisa diperoleh. Data tersebut bisa dimanfaatkan dan diolah menjadi informasi untuk meningkatkan penjualan produk. Seperti pola penjualan barang di toko sembako, hal inilah yang dijadikan sebagai dasar pengolahan data mining pada penjualan barang toko sembako. Dengan adanya kegiatan penjualan setiap hari, data akan semakin bertambah banyak, data tersebut tidak hanya berfungsi sebagai arsip bagi pengelola, data tersebut dapat diolah menjadi informasi yang bermanfaat untuk peningkatan penjualan barang di toko sembako. Pada dasarnya semua data memiliki informasi-informasi yang bermanfaat, yang bisa digunakan untuk mengambil suatu keputusan dan untuk memperoleh pengetahuan baru tentang pola penjualan barang elektronik.

Data mining merupakan salah satu cabang ilmu komputer yang banyak digunakan dan dipelajari oleh para ahli ilmu komputer dan programmer. Data mining merupakan serangkaian proses untuk menggali nilai tambah yang berupa informasi dari suatu database dengan melakukan pencarian pola-pola dari data, dengan tujuan memanipulasi data menjadi informasi yang lebih berharga. Data Mining merupakan serangkaian proses untuk mendapatkan nilai tambah dari kumpulan data. Kata mining sendiri artinya adalah usaha untuk mendapatkan nilai dan informasi dari sejumlah besar material dasar yang dimana disini adalah data. Karena itu Data Mining sebenarnya merupakan akar dari bidang ilmu database, statistik, dan machine learning.

Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Apriori dapat diimplementasikan pada sistem penjualan, dengan aplikasi berbasis teknologi informasi dihasilkan sebuah metode yang bisa meningkatkan penjualan dengan cara memberikan saran kepada konsumen dan keterkaitan barang yang dibeli oleh konsumen bisa dihitung dengan teknik algoritma apriori.

II. KAJIAN TEORI

A. Data Mining

Secara sederhana data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar [1]. Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data [2]. Data mining sering 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 [3].

Data mining merupakan suatu proses penambangan informasi penting dari suatu data. Teknik-teknik rumit tersebut nantinya akan mengidentifikasi dan mengekstraksi informasi yang bermanfaat dari suatu database besar. Data mining adalah berkaitan dengan bidang ilmu computer lain, misalnya seperti database system, statistik, komputasi tingkat tinggi, information retrieval, dan machine learning. Selain itu, data mining didukung oleh ilmu lain seperti neural network, pengenalan pola, spatial data analysis, image database, signal processing. Data mining didefinisikan sebagai proses untuk menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemuka harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar [4].

B. Tahap Data Mining

Sebagai rangkaian proses, data mining dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif, pengguna akan terlibat langsung atau dengan perantara knowledge base [5]. Tahap-tahap data mining ada 6 yaitu:

a. Pembersihan data (data cleaning)

Pembersihan data adalah proses untuk menghilangkan noise dan data yang tidak relevan. Pada umumnya data dari database perusahaan maupun hasil eksperimen sendiri,

memiliki isian yang tidak valid. Ada juga atribut data yang tidak sesuai dengan perkiraan data mining yang dimiliki. Data-data yang tidak sesuai itu akan dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data mining karena data yang digunakan akan berkurang kompleksitasnya.

b. Integrasi data (data integration)

Integrasi data merupakan penggabungan data ke dalam database baru. Tidak jarang ada data yang diperlukan tidak hanya berasal dari satu database tetapi juga beberapa dari database yang berbeda. Integrasi data dilakukan untuk atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut jenis produk, nomor pelanggan, nama dan lainnya. Kesalahan pada integrasi data dapat menghasilkan hasil yang berbeda dan menyimpang pengambilan aksi nantinya jadi harus dilakukan dengan cermat dan teliti.

c. Seleksi data (data selection)

Data yang akan digunakan di dalam database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Contohnya, sebuah kasus yang menilite factor kecenderungan manusia dalam membeli dalam market basket analysis, tidak perlu mengambil nama pelanggan, cukup hanya dengan id pelanggan saja.

d. Transformasi data (data transformation)

Data akan digabung atau diubah dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan data yang khusus sebelum bisa digunakan. Sebagai contoh metode seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karena itu data berupa angka yang berlanjut perlu dibagi menjadi beberapa interval.

e. Proses mining

Proses mining merupakan proses utama saat metode diterapkan untuk menemukan hasil yang berharga dan tersembunyi dari data tersebut.

f. Evaluasi pola (pattern evaluation)

Perlunya mengidentifikasi pola-pola yang menarik kedalam knowledge based berharga yang ditemukan. Dalam tahap ini hasil dari data mining berupa pola-pola yang khas maupun prediksi yang dievaluasi untuk menilai apakah perkiraan yang ada memang tercapai atau berhasil. Bila hasil yang didapat tidak sesuai hipotesa ada beberapa alternatif, seperti mencoba metode data mining lain yang lebih sesuai

C. Algoritma Apriori

Algoritma apriori adalah merupakan algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk penentuan frequent itemsets untuk aturan asosiasi boolean. Algoritma apriori termasuk jenis aturan asosiasi antara beberapa atribut yang sering disebut affinity analysis atau market basket analysis.

Algoritma apriori digunakan agar komputer dapat mempelajari aturan asosiasi, mencari pola hubungan antar satu atau lebih item dalam suatu dataset. Algoritma apriori banyak digunakan pada data transaksi [3]. Ada beberapa algoritma yang sudah dikembangkan mengenai aturan asosiasi, namun ada satu algoritma klasik yang masih tetap sering dipakai yaitu algoritma apriori. Ide dasar dari algoritma apriori adalah dengan mengembangkan frequent itemset. Secara rekursif mengembangkan frequent itemset dengan dua item, tiga item dan seterusnya hingga frequent itemset dengan semua ukuran. Untuk mengembangkan frequent set dengan dua item, dapat menggunakan frequent set item. Alasannya apabila set satu item tidak melebihi support minimum, maka sembarang ukuran itemset yang lebih besar tidak akan melebihi support minimum yang ada. Setiap tahap memerlukan sekali pemeriksaan ke seluruh isi database.

Dalam asosiasi terdapat antecedent dan consequent, antecedent untuk mewakili bagian “jika” dan consequent untuk mewakili “maka”. Dalam analisis ini, antecedent dan consequent adalah sekelompok item yang tidak punya hubungan secara Bersama [3].

Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, support (nilai penunjang) yaitu presentase kombinasi item tersebut dalam database dan confidence (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiatif. Analisis asosiasi didefinisikan suatu proses untuk menemukan semua aturan asosiatif yang memenuhi syarat minimum untuk support dan syarat minimum untuk confidence. Langkah pertama algoritma apriori adalah support dari setiap item dihitung dengan mengamati database. Setelah support dari setiap item didapat, item yang memiliki support lebih besar dari minimum support dipilih sebagai pola frekuensi tinggi dengan Panjang 1 atau 1-itemset. Kapanjangan kitemset yang berarti satu set yang terdiri dari k item.

Langkah kedua adalah menghasilkan 2-itemset yang tiap set-nya memiliki dua item. Pertama dibuat kandidat 2-itemset dari kombinasi semua 1-itemset. Lalu untuk tiap calon 2-itemset ini dihitung support dengan mengamati database. Support itu sendiri adalah jumlah transaksi dalam database yang mengandung kedua item dalam calon 2-itemset. Setelah support dari semua calon 2-itemset didapatkan, calon 2-itemset yang memenuhi syarat minimum support dapat ditetapkan sebagai 2-itemset yang berarti merupakan frekuensi tinggi dengan Panjang 2 [2]. Selanjutnya, semua iterasi ke-k dapat dibagi lagi menjadi beberapa bagian :

a. Kandidat k-itemset dibuat dari kombinasi k-itemset yang didapat dari iterasi sebelumnya. Salah satu ciri dari algoritma apriori adalah adanya pemangkasan kandidat k-itemset yang subset-nya berisi k-1 item tidak termasuk dalam pola frekuensi tinggi dengan Panjang k-1. Dari sini akan terbentuk kandidat itemset.

b. Support dari tiap kandidat k-itemset didapat dengan mengamati database untuk menghitung semua jumlah transaksi yang memuat semua item di dalam kandidat k-itemset. Ini ciri dari algoritma apriori yang dimana diperlukan penghitungan dengan mengamati seluruh database sebanyak k-itemset terpanjang.

c. Tetapkan pola frekuensi tinggi.

Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang dimana support lebih besar dari minimum support. Kemudian dihitung confidence masing-masing kombinasi item. Iterasi akan berhenti jika semua item telah dihitung sampai tidak ada kombinasi item lagi [2].

D. Asosiasi

Asosiasi adalah salah satu task data mining yang bertujuan untuk menemukan aturan asosiasif antara item-item data. Langkah utama yang digunakan dalam asosiasi yaitu mengetahui seberapa sering item combination muncul pada suatu database, yang disebut frequent pattern. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan 2 parameter yaitu support, item combination, dan confidence yang merupakan kuatnya hubungan antara suatu item dalam aturan asosiatif. Jika support itemset dari itemset 1 memenuhi minimum support threshold yang sudah ditentukan maka 1 adalah frequent k-itemset.

Secara umum frequent k-itemset dilambangkan dengan Lk. Sehingga dari hasil tersebut dapat disimpulkan bahwa nilai dari confidence yaitu persentase antara support (A U B) dengan support (A). Dimana A merupakan itemset utama sedangkan B adalah itemset yang dijadikan rekomendasi.

$$\text{Supports } (S) = |U| |T| \times 100\%$$

Rumus di atas merupakan rumus untuk nilai support setiap benda, dimana rumus ini akan digunakan sebagai pedoman barang mana saja yang akan menjadi yang terjual secara bersamaan dalam satu transaksi dan banyak terjual. U pada rumus adalah himpunan transaksi yang mengandung itemset S, sedangkan |U| merupakan jumlah elemen pada U dan |T| adalah jumlah transaksi yang terjadi.

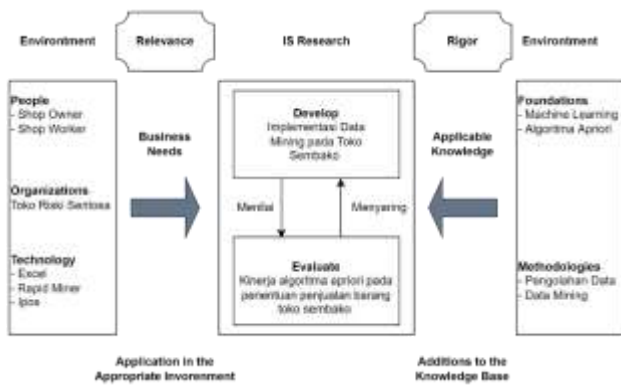
Confidence pada asosiasi adalah perbandingan antara nilai support dari himpunan item yang mendahuluinya. Contohnya adalah apabila ada “Jika X dan Y maka Z”, maka confidence dari asosiasi tersebut adalah :

$$\text{Confidence } (R) = \frac{\text{Support}\{X, Y, Z\}}{\text{Support}\{X, Y\}} \times 100\%$$

Dari rumus diatas bisa disimpulkan bahwa confidence dan support sangat berguna dalam penentuan apakah itemset memiliki keterkaitan yang efektif atau tidak.

III. METODE

Model Konseptual adalah pengembangan model yang bersifat analitis yang bertujuan untuk menguraikan dan menjelaskan variable yang digunakan dalam penelitian beserta keterkaitannya yang bertujuan untuk memetakan masalah serta kebutuhan yang diinginkan.



GAMBAR 1 Metodologi Konseptual

Pada Gambar III.1 dalam model konseptual merupakan penelitian yang digunakan untuk mengembangkan analisis untuk penjualan barang di toko sembako Riski Sentosa yang dilandasi oleh permasalahan shop owner dan shop worker yang berjualan sembako. Dengan itu penelitian ini bertujuan untuk memberikan solusi dari permasalahan tersebut, dengan membantu menganalisis barang yang terjual yang berfungsi sebagai alat bantu pengambilan keputusan dari shop owner dalam strategi meningkatkan pemasaran.

IV. HASIL DAN PEMBAHASAN

A. Proses Algoritma Apriori

Data diambil dari laporan penjualan disimpan dalam bentuk file excel atau *.xlsx atau *.xls. Data transaksi lalu dibersihkan dengan cara menghapus atribut-atribut yang tidak diperlukan pada penelitian yang dilakukan. Data yang sudah dibersihkan yaitu sebagai berikut.

TABEL 1 Data penjualan setelah dibersihkan

| No | i1 | i2 | i3 | i4 | i5 | i6 |
|----|----|----|----|----|----|----|
| 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 0 | 1 | 0 | 1 | 0 | 1 |
| 4 | 0 | 0 | 1 | 0 | 0 | 0 |
| 5 | 0 | 1 | 1 | 0 | 1 | 0 |
| 6 | 0 | 1 | 0 | 1 | 1 | 1 |
| 7 | 1 | 1 | 0 | 0 | 0 | 1 |
| 8 | 0 | 0 | 1 | 0 | 1 | 1 |
| 9 | 1 | 1 | 1 | 0 | 0 | 0 |
| 10 | 0 | 1 | 0 | 1 | 0 | 1 |
| 11 | 0 | 1 | 1 | 0 | 1 | 0 |
| 12 | 0 | 0 | 1 | 1 | 1 | 1 |
| 13 | 1 | 1 | 0 | 1 | 1 | 1 |
| 14 | 1 | 0 | 1 | 1 | 1 | 0 |
| 15 | 1 | 0 | 0 | 1 | 1 | 0 |

Tabel 1 merupakan data yang sudah di bersihkan dan sudah diubah sesuai dengan kebutuhan penelitian selanjutnya. Atribut yang diperlukan yaitu Id_produk dan nomor_transaksi. I1, i2, i3, i4 adalah barang-barang yang ada di toko sembako Riski Sentosa. Sedangkan untuk nilai 0 dan 1 dalam tabel bisa diartikan 0 adalah barang yang

tidak terjual dalam sebuah transaksi dan 1 adalah barang yang terjual dalam sebuah transaksi.

Data yang digunakan pada penelitian ini adalah data sampel dari transaksi penjualan sembako. Untuk mempermudah melakukan perhitungan maka dilakukan inisialisasi untuk masing-masing item. Data inisialisasi produk dapat dilihat pada tabel berikut ini:

TABEL 2 Inisialisasi Data

| No. | ID Barang | Inisialisasi |
|-----|-----------|--------------|
| 1 | ID1 | I1 |
| 2 | ID2 | I2 |
| 3 | ID3 | I3 |
| 4 | ID4 | I4 |
| 5 | ID5 | I5 |

Produk yang dijual toko sembako ada banyak tipe-tipe sembako namun yang dijadikan sampel dalam proses perhitungan data mining hanya mengambil lima jenis produk atau id produk.

TABEL 3 Data Transaksi Penjualan 1 sampai 5

| Transaksi ke- | Jenis Barang yang Dibeli |
|---------------|--------------------------|
| 1 | I1,I3,I5 |
| 2 | I4 |
| 3 | I1,I3,I5 |
| 4 | I1,I4 |
| 5 | I1,I5 |

Diambil sampel lima Transaksi barang sembako dari lima ID Barang yaitu transaksi pertama I1,I3,I5 , transaksi kedua I4, transaksi ketiga I1,I3,I5 transaksi empat I1,I4, transaksi lima I1,I5.

TABEL 4 Bentuk Tabular Transaksi Penjualan dari 1 sampai 5

| No | I1 | I2 | I3 | I4 | I5 |
|----|----|----|----|----|----|
| 1 | 1 | 0 | 1 | 0 | 1 |
| 2 | 0 | 0 | 0 | 1 | 0 |
| 3 | 1 | 0 | 1 | 0 | 1 |
| 4 | 1 | 0 | 0 | 1 | 0 |
| 5 | 1 | 0 | 0 | 0 | 1 |

Setelah tabel dalam bentuk tabular selesai maka langkah selanjutnya yaitu penentuan frequent 1-itemset seperti terlihat dalam tabel di bawah ini:

TABEL 5 Perhitungan Frequent 1 - Itemset

| No. | Item | Support |
|-----|------|--------------------|
| 1 | I1 | (4/5) x 100% = 80% |
| 2 | I2 | (0/5) x 100% = 0% |
| 3 | I3 | (2/5) x 100% = 40% |
| 4 | I4 | (2/5) x 100% = 40% |
| 5 | I5 | (3/5) x 100% = 60% |

Data di atas menggambarkan bentuk data 1-itemset, yang terdiri atas attribute item sebagai nama item yang ada. Dari Tabel perhitungan frequent 1-itemset di atas,

ditetapkan nilai minimum support sebesar 0,25 (25%). Maka item yang memenuhi nilai minimum support 0.25 (25%) adalah sebagai berikut:

TABEL 6 Frequent 1 – Itemset Memenuhi Minimum Support

| No. | Item | Support |
|-----|------|-----------------------------|
| 1 | I1 | $(4/5) \times 100\% = 80\%$ |
| 2 | I3 | $(2/5) \times 100\% = 40\%$ |
| 3 | I4 | $(2/5) \times 100\% = 40\%$ |
| 4 | I5 | $(3/5) \times 100\% = 60\%$ |

Data di atas merupakan frequent 1-itemset yang memenuhi nilai minimum support. Data tersebut akan digunakan untuk membentuk pola kombinasi mulai dari pembentukan pola kombinasi 2- itemset sampai pada pembentukan pola kombinasi 3- itemset. Pembentukan frequent 2-itemset dibentuk dengan mengkombinasikan item-item hasil perhitungan frequent 1-itemset yang memenuhi nilai minimum support yang telah kita tentukan. Kombinasi yang dibentuk dari frequent 2- itemset dapat dilihat pada tabel di bawah ini:

TABEL 7 Perhitungan Frequent 2 – Itemset

| No. | Item | Support |
|-----|--------|-----------------------------|
| 1 | I1, I3 | $(2/5) \times 100\% = 40\%$ |
| 2 | I1, I5 | $(3/5) \times 100\% = 60\%$ |
| 3 | I3, I5 | $(2/5) \times 100\% = 40\%$ |
| 4 | I1, I4 | $(0/5) \times 100\% = 0\%$ |
| 5 | I3, I4 | $(0/5) \times 100\% = 0\%$ |
| 6 | I4, I5 | $(0/5) \times 100\% = 0\%$ |

Data di atas menggambarkan bentuk data kombinasi 2-itemset. Dari Tabel perhitungan frequent 2-itemset di atas, ditetapkan nilai minimum support sebesar 0,25 (25%). Maka item yang memenuhi nilai minimum support 0.25 (25%) adalah sebagai berikut:

TABEL 8 Frequent 2 – Itemset yang Memenuhi Minimum Support

| No. | Item | Support |
|-----|--------|-----------------------------|
| 1 | I1, I3 | $(2/5) \times 100\% = 40\%$ |
| 2 | I1, I5 | $(3/5) \times 100\% = 60\%$ |
| 3 | I3, I5 | $(2/5) \times 100\% = 40\%$ |

dapatkan, maka langkah selanjutnya yaitu melakukan pembentukan kombinasi frequent 3- itemset. Pembentukan kombinasi frequent 3- itemset dilakukan dengan cara menggabungkan item I1, I3, I5. Kombinasi yang dibentuk dari frequent 3-itemset dapat dilihat pada tabel di bawah ini:

TABEL 9 Perhitungan Frequent 3 – Itemset

| No. | Item | Support |
|-----|------------|-----------------------------|
| 1 | I1, I3, I5 | $(2/5) \times 100\% = 40\%$ |

Dari tabel di atas tidak ada nilai yang memenuhi minimum support 0.25 (25%), sehingga perhitungan frequent itemset dihentikan.

B. Proses dan Hasil Algoritma Apriori



GAMBAR 3 Proses Algoritma Apriori

Gambar di atas adalah proses algoritma apriori di Rapid Miner. Pada algoritma apriori ini juga terdiri dari beberapa tahapan yaitu pertama adalah menginputkan data bersih dan siap untuk diolah. Tahapan kedua juga mengubah bilangan numerical menjadi binominal, tahapan ini bertujuan supaya data dapat dimengerti oleh komputer. Tahapan terakhir adalah memasukkan W-Apriori yang berarti memasukkan algoritma apriori. Pada tahapan ini nilai support adalah 0,25 dan nilai confidence adalah 0,5. Sehingga yang didapatkan dari proses algoritma apriori adalah sebagai berikut:

```

Minimum support: 0.25 (582 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 15

Generated sets of large itemsets:

Size of set of large itemsets L(1): 178
Size of set of large itemsets L(2): 1435

Best rules found:

1. i44=true i163 ==> i142=true 632   conf:(0.54)
2. i142=true i164 ==> i44=true 632   conf:(0.54)
3. i10=true i137 ==> i173=false 614   conf:(0.54)
4. i92=false i146 ==> i173=false 615   conf:(0.54)
5. i23=true i119 ==> i59=true 597   conf:(0.53)
6. i121=true i140 ==> i131=true 608   conf:(0.53)
7. i53=false i125 ==> i136=true 600   conf:(0.53)
8. i5=true i135 ==> i21=true 605   conf:(0.53)
9. i132=false i136 ==> i131=true 605   conf:(0.53)
10. i152=true i141 ==> i100=false 606   conf:(0.53)
    
```

GAMBAR 2 Hasil Algoritma Apriori

Dari hasil algoritma apriori, terdapat 10 kombinasi terbaik dari proses yang telah dilakukan. 178 itemset pertama untuk kombinasi atau pola yang terbuat, lalu ada 1435 itemset kedua untuk kombinasi atau pola yang terbuat. Dari Gambar 2, hasil pada *rules* pertama bahwa $i44 = true \ i163 \rightarrow i142 = true \ 632 \ conf: (0.54)$ yang artinya untuk barang $i44$ memiliki jumlah total sebanyak 1163 dari jumlah data yang digunakan sebanyak 2328 data pada proses algoritma apriori dan memiliki hubungan asosiasi dengan barang $i142$ sebanyak 632, sehingga memiliki nilai confidence sebesar 0,54. Nilai *confidence* yang didapatkan memiliki arti kombinasi atau pola yang terbuat adalah hubungan kombinasi yang pasti terjadi. Begitu pula untuk *rules* yang ada selanjutnya dalam hasil algoritma apriori.

Selanjutnya menguji tingkat validitas dengan melakukan perhitungan nilai *lift* antara kombinasi item yang sudah terbentuk. Pengujian ini bertujuan untuk melihat seberapa pola kombinasi item yang terbuat akurat, dengan melihat apakah nilai *lift* lebih dari nilai 1. Jika nilai *lift* sama dengan nilai 1 atau lebih, maka kombinasi item yang terbentuk adalah valid. Namun jika nilai *lift* kurang dari nilai 1, berarti kombinasi item yang terbuat kurang akurat dan tidak valid. Rumus nilai *lift* yang digunakan adalah:

$$Lift = \frac{Support\ A\ dan\ B}{Support\ (A)\ X\ Support\ (B)}$$

Berdasarkan rumus *lift* tersebut, berikut ini contoh data hasil dari kombinasi item dari perhitungan yang dilakukan:

TABEL 10 Nilai Lift

| Premises | Conclusion | Lift |
|----------|------------|------|
| I42 | I62 | 1.08 |
| I8 | I54 | 1.07 |

Pada Tabel V.1 item I42 dan I62 memiliki nilai *lift* sebesar 1.08 yang berarti nilai kombinasi item tersebut sudah akurat dan valid begitu pula dengan kombinasi item I8 dan I54.

C. Analisis Hasil

Dari hasil uji algoritma apriori didapatkan 30 kombinasi item terbaik yang sudah membentuk pola hubungan karena telah memenuhi aturan asosiasi yang telah ditentukan yaitu dengan nilai *support* dan nilai *confidence*, hanya ada 7 kombinasi item yang dimana antara item menunjukkan *true* dan *true* yang artinya kedua barang dibeli secara bersamaan. Table berikut adalah 7 kombinasi item tersebut.

| No. | Item 1 | Item 2 | Lift |
|-----|--------|--------|------|
| 1 | I44 | I142 | 1.09 |
| 2 | I42 | I62 | 1.08 |
| 3 | I74 | I154 | 1.08 |
| 4 | I8 | I54 | 1.07 |
| 5 | I5 | I21 | 1.07 |
| 6 | I70 | I175 | 1.07 |
| 7 | I23 | I59 | 1.07 |

TABEL 3 Hasil Uji Algoritma Apriori

Dari hasil uji algoritma apriori pada Tabel V.2 setiap satu kombinasi item memiliki kemungkinan dibeli satu sama lain. Dengan ini rekomendasi kombinasi item

tersebut dapat dijadikan strategi pemasaran dari Toko Riski Sentosa untuk pelanggan baru. Hasil penelitian ini dapat digunakan untuk penempatan barang, dengan menempatkan barang sesuai dengan table kombinasi item diatas, agar memudahkan pelanggan untuk membeli dua barang. Rekomendasi juga bisa berupa dengan menggunakan paket promo antara kombinasi item atau dengan memberikan diskon jika membeli 2 kombinasi item tersebut.

V. KESIMPULAN

Penelitian ini dilakukan untuk mencari keterkaitan antara barang yang terjual di Toko Sembako Riski Sentosa. Berdasarkan penelitian yang dilakukan, kesimpulan yang dapat diambil adalah algoritma apriori adalah salah satu algoritma yang digunakan dalam memprediksi kedekatan hubungan suatu item dalam jumlah data yang besar. Perhitungan nilai *support* dan nilai *confidence* sangat mempengaruhi dalam penentuan kombinasi item yang terbentuk. Pada penelitian ini nilai *support* yang ditetapkan adalah 0.25%. sedangkan untuk nilai *confidence* yang didapatkan adalah 0.5%. Semakin besar nilai *support* dan nilai *confidence* yang didapatkan maka hasil yang diperoleh semakin maksimal. Lalu juga terdapat perhitungan nilai *lift*, nilai *lift* ini digunakan untuk melihat tingkat keakuratan kombinasi yang terbentuk dengan kombinasi item tersebut harus bernilai *lift* sama dengan atau lebih dari 1. Maka dari perhitungan tersebut didapat 7 kombinasi item yang memenuhi aturan tersebut dan kombinasi barang yang paling banyak terjual dalam 1 bulan adalah i44 dan i142.

REFERENSI

- [1] Davies, and Paul Beynon, Database Systems Third Edition, New York: Palgrave Macmillan, 2004.
- [2] Pramudiono, I, Pengantar Data Mining : Menambang Permata, Surabaya: Paper ITS, 2007.
- [3] B. Santoso, Pengaruh Harga, Promosi, Kualitas Layanan terhadap Kepuasan Pelanggan Jasa Ojek Online, Jember: Jurnal Sains Manajemen dan Bisnis Indonesia, 2017.
- [4] Ian H.Witten, E. F., Data mining: Practical Machine Learning Tools and Techniques, Second Edition, San Fransisco: Elsevier Inc, 2005.
- [5] Kusriani dan Emha Taufiq Luthfi, Algoritma Data Mining, Yogyakarta: Andi, 2010.