
PENERAPAN ALGORITMA APRIORI PADA TRANSAKSI PENJUALAN AIR MINUM UNTUK MENINGKATKAN STRATEGI BISNIS (STUDI KASUS: PT SILA TIRTA GEMILANG)

Mayland Trifena¹⁾, Nono Heryana²⁾, Taufik Ridwan³⁾

^{1,2,3}Prodi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

Correspondence author: M. Trifena, 2010631250060@student.unsika.ac.id, Karawang, Indonesia

Abstract

In the era of business growth supported by information technology, business competition and information needs are increasing. However, many businesses, including PT Sila Tirta Gemilang, have not utilized it optimally. The company faces challenges in managing drinking water stock and predicting sales efficiently, which leads to stock buildup and demand uncertainty. This research aims to apply data mining to PT Sila Tirta Gemilang's sales strategy in the past year. The research uses the Knowledge Discovery in Databases (KDD) method with six stages: data selection, data cleaning, data transformation, data mining, evaluation, and knowledge presentation. The apriori algorithm is used to determine the frequency of item sets and find customer purchase patterns. The types of drinking water studied include D. 200 ML, D. 600 ML, K. 200 ML, K. 600 ML, S. 200 ML, QUA. F, Q. 600 ML, COCO, Fresh Tea, and GMES. The results showed that the largest support value for one item was 62.61%, two items were 24.9%, and three items were 7.2%. Overall, the confidence value is 89.5%, and the lift ratio is 1.670. The resulting 55 association rules can be used by companies to improve sales efficiency and effectiveness.

Keywords: *stock management, sales prediction, data mining, sales strategy, apriori algorithm*

Abstrak

Dalam era pertumbuhan bisnis yang didukung oleh teknologi informasi, persaingan bisnis dan kebutuhan informasi semakin meningkat. Namun, banyak bisnis, termasuk PT Sila Tirta Gemilang yang bergerak dalam penjualan air minum, belum memanfaatkannya secara optimal. Perusahaan ini menghadapi tantangan dalam mengelola stok air minum dan memprediksi penjualan secara efisien yang menyebabkan penumpukan stok dan ketidakpastian permintaan. Penelitian ini bertujuan menerapkan *data mining* untuk strategi penjualan PT Sila Tirta Gemilang dalam satu tahun terakhir. Penelitian menggunakan metode Knowledge Discovery in Database (KDD) dengan enam tahapan, data selection, data cleaning, data transformation, data mining, evaluation, dan knowledge presentation. Algoritma apriori digunakan untuk mengetahui frekuensi itemset dan mencari pola pembelian pelanggan. Jenis air minum yang diteliti meliputi D. 200 ML, D. 600 ML, K. 200 ML, K. 600 ML, S. 200 ML, QUA. F, Q. 600 ML, COCO, Fresh Tea, dan GMES. Hasil penelitian menunjukkan nilai support terbesar untuk satu itemset adalah 62.61%, dua itemset 24.9%, dan tiga itemset

7.2%. Secara keseluruhan, nilai confidence sebesar 89.5% dan lift ratio 1.670. Dihasilkan 55 aturan asosiasi yang dapat digunakan perusahaan untuk meningkatkan efisiensi dan efektivitas penjualan.

Kata Kunci: pengelolaan stok, prediksi penjualan, *data mining*, strategi penjualan, algoritma apriori

A. PENDAHULUAN

Persaingan bisnis dan kebutuhan akan informasi semakin meningkat seiring dengan pertumbuhan bisnis dibantu dengan teknologi informasi. Dengan teknologi ini, para pelaku bisnis dapat mendapatkan informasi yang dibutuhkan untuk membangun solusi dan merancang strategi baru untuk memenuhi kebutuhan pelanggan yang semakin meningkat (Takdirillah, 2020).

Banyak bisnis masih gagal memanfaatkan teknologi informasi, khususnya komputasi dengan optimal. Bisnis biasanya menyimpan catatan transaksi penjualan setiap bulan, tetapi mereka hanya memberikan laporan setiap minggu dan setiap bulan. Jika data tidak ditangani, peningkatan data akan menghasilkan banyak data tetapi sedikit informasi. Akibatnya, hanya akumulasi data yang tidak menguntungkan akan terjadi (Riszky & Sadikin, 2019).

Beberapa perusahaan masih mengandalkan strategi pemasaran produk secara manual tanpa dasar yang jelas hanya berdasarkan perkiraan. Untuk mendukung penerapan strategi pemasaran yang lebih efektif, perusahaan seharusnya memanfaatkan kemajuan teknologi komputasi. Data transaksi penjualan yang terdokumentasi dapat menjadi alat yang sangat berharga bagi manajemen perusahaan dan membantu meningkatkan penjualan melalui promosi produk yang lebih terarah (Dora et al., 2023).

Salah satunya pada PT Sila Tirta Gemilang yang bergerak dalam penjualan air minum juga menghadapi tantangan serupa. PT Sila Tirta Gemilang menghadapi

tantangan dalam mengelola stok air minum secara efisien dan melakukan prediksi penjualan air minum yang tepat. Masalah yang kerap muncul melibatkan penumpukan barang yang berlebihan dalam gudang, kehabisan stok saat ada permintaan pelanggan, dan kurangnya penggunaan prediksi penjualan yang masih dilakukan secara manual.

Perusahaan hingga saat ini belum memiliki sistem yang mampu memprediksi penjualan air minum dengan akurat. Keputusan terkait manajemen stok dan perencanaan produksi masih bergantung pada pengalaman dan intuisi karyawan perusahaan. Penerapan teknologi dan analisis data menjadi solusi penting dalam mengelola stok dan memprediksi penjualan (Wulandhari et al., 2023).

Salah satu pendekatan yang dapat digunakan adalah *data mining* dengan menggunakan algoritma apriori. Melalui penerapan data mining dan algoritma Apriori, diharapkan perusahaan dapat menganalisis data penjualan sebelumnya untuk mengidentifikasi pola-pola yang dapat digunakan dalam memprediksi permintaan pelanggan di masa depan. Dengan demikian, perusahaan dapat merencanakan produksi dan stok dengan lebih akurat, mengurangi pemborosan, dan menghindari ketidaktersediaan barang.

Pada umumnya memang banyak faktor yang mempengaruhi ketidakpastian dalam penjualan maka perlu adanya analisis penjualan lebih lanjut yang dapat membantu manajemen toko dalam mengoptimalkan pembelian produk dan menentukan strategi penjualan berdasarkan kondisi pasar, salah satunya adalah dengan pemanfaatan data transaksi. Oleh karena

itu, penelitian ini akan menggunakan metode asosiasi untuk menganalisis pola penjualan air minum di PT Sila Tirta Gemilang dari data transaksi penjualan. Asosiasi adalah metode yang digunakan untuk menemukan pola atau informasi yang berguna dari data yang besar dan kompleks (Ikhwan & Aslami, 2020).

Dalam hal ini, penelitian akan memanfaatkan data historis penjualan air minum selama satu tahun untuk mengidentifikasi pola-pola pembelian yang signifikan. Dengan memanfaatkan pola-pola tersebut, diharapkan PT Sila Tirta Gemilang dapat meningkatkan strategi bisnis dalam penjualan. Ini akan membantu meningkatkan efisiensi dalam pengelolaan stok air minum dengan menentukan jenis produk air minum yang perlu diproduksi dan dijaga stoknya dengan prioritas tertentu.

Algoritma Apriori dipilih dalam penelitian ini berdasarkan hasil positif dari penelitian sebelumnya, seperti penelitian yang dilakukan oleh Addzifi Mochamad Gumelar dan rekannya dimana mereka menggunakan Algoritma Apriori dalam data mining untuk menentukan korelasi produk berdasarkan pola transaksi penjualan. Penelitian tersebut menunjukkan bahwa distributor dapat mengidentifikasi produk yang sering dibeli melalui sistem ini. Pemimpin bisnis dapat memanfaatkan informasi tersebut untuk meningkatkan strategi penjualan dan pemasaran. Selain itu, sistem ini memberikan manfaat bagi pembeli dengan menampilkan hubungan produk yang sering dibeli bersamaan. Dengan menggunakan metode asosiasi dan algoritma apriori, sistem ini dapat mengikuti dan mengontrol permintaan pasar secara efektif (Gumelar & Akbar, 2020).

Selanjutnya penelitian yang dilakukan oleh Nita Syahputri dalam penelitiannya menggunakan algoritma apriori memberikan kemudahan implementasi, proses yang singkat, fleksibilitas, dan sederhana dalam prinsip yang dijelaskan

secara non-statistik. Dari penelitian tersebut, ditemukan 142 aturan asosiasi dengan support di atas 10% dan *confidence* 50%, menunjukkan bahwa metode apriori dapat diterapkan secara efektif pada data sampel penjualan. Hasil penelitian ini diharapkan memberikan bantuan bagi konsumen dan pemilik usaha dalam menentukan makanan serta minuman yang paling diminati (Syahputri, 2020).

Penelitian yang dilakukan oleh Syukron Anas dan rekan-rekannya, mengenai perbandingan penelitian menggunakan algoritma apriori dan FP-Growth didapatkan bawah dari hasil kedua algoritma ini sama-sama menghasilkan dua aturan, namun hasil yang ditunjukkan oleh algoritma FP-Growth mengalami beberapa kali kesamaan barang kombinasi, sedangkan Algoritma Apriori memiliki lebih banyak variasi. Hasil yang didapatkan algoritma apriori mempunyai nilai support dan *confidence* bernilai 98,4 sedangkan algoritma FP-Growth mempunyai nilai support dan *confidence* bernilai 95,2 (Anas et al., 2022).

Berdasarkan penelitian sebelumnya dan perbandingan algoritma dapat disimpulkan bahwa penggunaan algoritma apriori menggunakan metode asosiasi dalam menganalisis pola penjualan air minum di PT Sila Tirta Gemilang dapat membantu perusahaan dalam merencanakan produksi air minum dengan lebih efektif dan efisien, mengurangi penumpukan barang di gudang, serta meningkatkan keuntungan dan mengurangi biaya produksi maupun penjualan.

Dengan demikian, akan digunakan Algoritma Apriori untuk mengekstraksi pola-pola penjualan air minum dalam memproduksi produk yang relevan. Selanjutnya, pola-pola ini akan diolah menggunakan metode asosiasi untuk mendapatkan informasi yang berguna bagi PT Sila Tirta Gemilang dalam menentukan jumlah stok air minum yang dibutuhkan. Metode asosiasi digunakan untuk memproses data yang kompleks dan besar

sehingga menghasilkan informasi yang berguna dan mudah dipahami oleh pengguna.

B. METODE PENELITIAN

Penelitian ini menggunakan metodologi *Knowledge Discovery in Databases* (KDD) dengan algoritma apriori, dipilih karena kecepatan dan ketepatannya dalam menemukan aturan asosiasi. KDD efektif mengidentifikasi pola dan tren dalam data, cocok untuk penelitian yang bertujuan menemukan pola penjualan atau asosiasi produk dalam dataset transaksi penjualan air minum. KDD mengungkap informasi berharga dan pola tersembunyi dalam database besar yang sebelumnya tidak terdeteksi tetapi bermanfaat (Erwansyah et al., 2021). *Data mining*, sebagai tahap dalam proses iteratif KDD, digunakan untuk mengekstrak pola dan hubungan yang memberikan wawasan berharga dari data tersebut

1. Pengumpulan Data

Pada tahapan pertama yaitu pengumpulan data dimana dilakukan dengan mengambil data dari perusahaan yang nantinya diolah menggunakan metodologi KDD.

2. *Data Selection*

Fase ini melibatkan transformasi data yang tidak memiliki entitas jelas menjadi data yang siap untuk proses data mining. Misalnya, dalam dataset penjualan produk yang besar, jika hanya tertarik pada informasi penjualan bulan terakhir, pemilihan data akan mengekstrak hanya data penjualan dari bulan terakhir untuk analisis lebih lanjut.

3. *Processing Data/Data Cleaning*

Processing Data/Data cleaning adalah proses mengeliminasi elemen yang tidak diinginkan dari dataset yang tidak konsisten atau tidak relevan. Ini mencakup penanganan data yang hilang, data tidak valid, atau kesalahan

ketik (Kurnia et al., 2020). Data yang dianggap tidak esensial dapat dihapus, suatu tindakan yang dikenal sebagai *data cleaning*. Proses ini sangat penting karena hasil teknik data mining dapat dipengaruhi oleh kualitas *data cleaning*.

4. *Data Transformation*

Proses mengubah dan menggabungkan data ke dalam bentuk tertentu disebut transformasi data. *Data mining* memerlukan format data tertentu; misalnya, metode seperti analisis asosiasi dan *clustering* hanya menerima data kategorial. Oleh karena itu, data yang berupa angka numerik yang memiliki sifat continue harus dibagi menjadi beberapa interval.

5. *Data Mining*

Penambangan data yang juga dapat disebut sebagai proses data mining merupakan tahapan kunci yang menggunakan metode untuk mengungkapkan pengetahuan berharga yang tersembunyi dalam data. Pada penelitian ini menggunakan aturan Algoritma Apriori. Algoritma Apriori dalam data mining digunakan untuk mengidentifikasi frekuensi dan hubungan antara itemset, berdasarkan nilai minimum *support* dan *confidence* yang telah ditetapkan (Firdaus et al., 2021).

6. *Evaluation*

Tahap evaluasi dalam KDD menilai hasil analisis dan pola yang ditemukan untuk menentukan nilai dan keberhasilan dalam mencapai tujuan bisnis atau penelitian. Proses ini melibatkan pengukuran kualitas dan relevansi pengetahuan yang ditemukan, serta menilai sejauh mana hasil tersebut memberikan wawasan berharga atau berkontribusi terhadap pemecahan masalah.

7. *Knowledge*

Tahap di mana informasi yang dihasilkan dari analisis data yang dapat memberikan wawasan atau pemahaman baru yang berguna.

C. HASIL DAN PEMBAHASAN

Tahapan penelitian ini akan menyajikan analisis mengenai kebutuhan yang diperlukan untuk memperoleh data yang relevan, tools, algoritma, dan metode KDD yang digunakan dalam penelitian ini.

1. Pengumpulan Data

Penelitian ini menggunakan data yang dikumpulkan dari PT Sila Tirta Gemilang. Dimana data tersedia dalam format *Microsoft Excel* dan mencakup catatan penjualan dari 25 Oktober 2022 hingga 31 Oktober 2023. Dengan total 5.808 catatan transaksi, data yang terorganisir ini memudahkan penulis dalam menjalankan penelitian.

2. Data Selection

Tahap ini bertujuan menyaring data untuk mempermudah proses berikutnya dengan mengidentifikasi fitur yang relevan. Meskipun basis data transaksi penjualan air minum memiliki banyak atribut, tidak semuanya diperlukan untuk analisis. Oleh karena itu, seleksi data dilakukan untuk memilih fitur yang penting dan menghapus yang tidak relevan. Pemodelan data mining dengan algoritma apriori akan menggunakan fitur-fitur terpilih. Tiga fitur yang dipilih adalah Tanggal, Keterangan, dan Produk, seperti ditunjukkan dalam tabel 1. Fitur-fitur ini menjadi fokus penelitian karena kombinasi nilai-nilainya menunjukkan pola pembelian konsumen pada hari yang sama. Dalam tahap transformasi data, nilai-nilai fitur produk akan disusun dalam tabel data.

Tabel 1. Hasil *Data Selection*

| No | Tanggal | Keterangan | Produk |
|----|------------|---------------------------|-----------------------|
| 1 | 25/10/2022 | TK BP BAMBANG PUTRI | D.220 ML, S.200 ML |
| 2 | 25/10/2022 | TK HR | D.220 ML |
| 3 | 25/10/2022 | TK AA JAYA/ACEN G | D.220 ML, S.200 ML |

| No | Tanggal | Keterangan | Produk |
|-------|------------|------------------|-----------------------|
| 4 | 25/10/2022 | TK BINTANG | K.200 ML |
| 5 | 25/10/2022 | TK ROI GAS | D.220 ML, S.200 ML |
| 6 | 25/10/2022 | TK MPT | D.220 ML, K.200 ML |
| 7 | 25/10/2022 | TK MPT | K.220 ML, S.200 ML |
| 8 | 25/10/2022 | TK BATARA | K.220 ML |
| 9 | 25/10/2022 | TK BATARA | D.220 ML, K.220 ML |
| 10 | 26/10/2022 | TK EMBUN PAGI | D.220 ML, K.200 ML |
| ... | ... | ... | ... |
| 5.808 | 31/10/2023 | TK HASIBUAN | D.220 ML |

3. Data Cleaning

Pada periode 25 Oktober 2022 hingga 31 Oktober 2023, data transaksi mengalami masalah nilai yang hilang dan duplikasi per tanggal. Penghapusan duplikasi diperlukan untuk memastikan akurasi model dalam *data mining*. Data dengan duplikasi nama pelanggan akan disatukan dan dihilangkan. Hasil pembersihan data ditunjukkan dalam tabel 1 dan tabel 2, yang menggambarkan kondisi sebelum dan sesudah penghapusan duplikasi.

```
[120] df['Tanggal'] = pd.to_datetime(df['Tanggal'], format='%d/%m/%Y')
[121] df_grouped = df.groupby(['Tanggal', 'Nama Pelanggan'], as_index=False).sum()
[122] df_cleaned = df_grouped.drop_duplicates()
[123] df_cleaned.fillna(0, inplace=True)
df_cleaned.shape
df_cleaned
```

Gambar 1. Menghapus Duplikasi Data

```
df_cleaned.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4291 entries, 0 to 4290
Data columns (total 24 columns):
#   Column              Non-Null Count  Dtype
---  ---
0   Tanggal              4291 non-null   datetime64[ns]
1   Nama Pelanggan       4291 non-null   object
2   D.220 ML              4291 non-null   object
3   D.600 ML              4291 non-null   object
4   D.1500 ML             4291 non-null   float64

[137] df_cleaned.isnull().sum()
Tanggal              0
Nama Pelanggan       0
D.220 ML              0
D.600 ML              0
D.1500 ML             0
```

Gambar 2. Mengecek Missing Value

Setelah penghapusan, jumlah transaksi berkurang dari 5.808 menjadi 4.291 records.

Tabel 2. Hasil Data Cleaning

| No | Tanggal | Keterangan | Produk |
|-------|------------|---------------------|------------------------------|
| 1 | 25/10/2022 | TK BP BAMBANG PUTRI | D.220 ML, S.200 ML |
| 2 | 25/10/2022 | TK HR | D.220 ML |
| 3 | 25/10/2022 | TK AA JAYA/ACENG | D.220 ML, S.200 ML |
| 4 | 25/10/2022 | TK BINTANG | K.200 ML |
| 5 | 25/10/2022 | TK ROI GAS | D.220 ML, S.200 ML |
| 6 | 25/10/2022 | TK MPT | D.220 ML, K.200 ML, S.200 ML |
| 7 | 25/10/2022 | TK BATARA | D.220 ML, K.220 ML |
| 8 | 26/10/2022 | TK EMBUN PAGI | D.220 ML, K.200 ML |
| 9 | 26/10/2022 | TK LIJORA | K.200 ML |
| 10 | 26/10/2022 | TK HASIBUAN | D.220 ML, K.220 ML, S.200 ML |
| ... | ... | | |
| 4.291 | 31/10/2023 | TK HASIBUAN | D.220 ML |

4. Data Transformation

Pada tahap pembentukan data mining dengan *RapidMiner*, dataset harus berbentuk nominal. Data yang dikumpulkan sebelumnya masih berupa numerik, sehingga perlu dilakukan transformasi dari numerik ke nominal. Transformasi ini mempermudah pembacaan dan analisis data penjualan. Dalam proses ini, data diubah dari bentuk numerik menjadi nominal menggunakan angka 1 dan 0. Angka 1 menunjukkan produk terjual, sedangkan angka 0 menunjukkan produk tidak terjual.

```
[185] process_data = df_cleaned.drop(columns=['Nama Pelanggan', 'Tanggal'])
[186] df1 = process_data
[189] df1 = df1.applymap(lambda x: 1 if x != 0 else 0)
[190] df1
```

Gambar 3. Mengubah Data Numerik Menjadi Nominal

Hanya data item yang digunakan dalam transformasi ini, sementara data nama pelanggan dan tanggal dihapus, menghasilkan bentuk seperti yang ditunjukkan pada gambar 1.

Gambar 4. Hasil Data Transformation

5. Data Mining

Tahap data mining dimulai dengan memilih algoritma apriori untuk mengidentifikasi pola transaksi penjualan dan memberikan rekomendasi jenis air minum. Proses dilakukan dengan tools *RapidMiner* yang berfokus pada identifikasi frekuensi itemset yang memenuhi batas minimum *support* dan *confidence*. Batas minimum *support* dipilih 3% dan batas minimum *confidence* dipilih 60%.

a. Analisa Pola Frekuensi Tinggi

Tahap ini bertujuan menemukan kombinasi itemset yang memenuhi batas *minimum support*. Nilai *support* dihitung dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi A}}{\text{Total Transaksi}} \times 100\% \quad (1)$$

Tabel 3. Daftar 1 Item Support 3%

| No | Item | Support count | Support (%) |
|----|----------|---------------|-------------|
| 1 | ICE TEA | 136 | 3.17% |
| 2 | COCO | 219 | 5.10% |
| 3 | GMES | 311 | 7.24% |
| 4 | K.220 ML | 325 | 7.57% |
| 5 | D.600 ML | 336 | 7.82% |

| No | Item | Support count | Support (%) |
|----|-----------|---------------|-------------|
| 6 | K.600 ML | 611 | 14.23% |
| 7 | S.200 ML | 1187 | 27.66% |
| 8 | K. 200 ML | 2238 | 52.16% |
| 9 | D.220 ML | 2687 | 62.61% |

Tabel 4. Daftar 2-Itemset Support 3%

| No | Kombinasi Dua Item | | Support Count | Support (%) |
|----|--------------------|-----------|---------------|-------------|
| 1 | S. 200 ML | GMES | 128 | 3.0% |
| 2 | K. 200 ML | K. 220 ML | 133 | 3.1% |
| 3 | K. 200 ML | COCO | 138 | 3.2% |
| 4 | D. 220 ML | COCO | 140 | 3.3% |
| 5 | D. 220 ML | K. 220 ML | 173 | 4% |
| 6 | K. 200 ML | D. 600 ML | 178 | 4.1% |
| 7 | K. 200 ML | GMES | 177 | 4.1% |
| 8 | D. 220 ML | D. 600 ML | 217 | 5.1% |
| 9 | D. 220 ML | GMES | 244 | 5.7% |
| 10 | S. 200 ML | K. 600 ML | 254 | 5.9% |
| 11 | D. 220 ML | K. 600 ML | 316 | 7.4% |
| 12 | K. 200 ML | K. 600 ML | 433 | 10% |
| 13 | K. 200 ML | S. 200 ML | 599 | 14% |
| 14 | D. 220 ML | S. 200 ML | 654 | 15.2% |
| 15 | D.220 ML | K. 200 ML | 1067 | 24.9% |

Tabel 5. Daftar 3-Itemset Support 3%

| No | Kombinasi Tiga Item | | | Support count | Support (%) |
|----|---------------------|----------|----------|---------------|-------------|
| 1 | D.220 ML | K.200 ML | S.200 ML | 309 | 7.2% |
| 2 | D.220 ML | K.200 ML | K.600 ML | 206 | 4.8% |
| 3 | K.200 ML | S.200 ML | K.600 ML | 179 | 4.2% |
| 4 | D.220 ML | K.200 ML | GMES | 140 | 3.3% |
| 5 | D.220 ML | S.200 ML | K.600 ML | 132 | 3.1% |

b. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, langkah berikutnya adalah mencari aturan asosiasi yang memenuhi syarat *minimum confidence*. Untuk menghitung *confidence* aturan asosiasi $A \rightarrow B$, digunakan persamaan berikut:

$$Confidence P(B|A) = \frac{\text{Jumlah Transaksi AUB}}{\text{Jumlah Transaksi A}} \times 100\% \quad (2)$$

Tabel 6. Hasil Nilai Minimum Confidence

| No | Rule | Support | Confidence (%) |
|-----|---|---------|----------------|
| 1 | Jika membeli jenis air minum KAS/ 212 maka akan membeli jenis air minum K. 200 ML | 76/124 | 61.3% |
| 2 | Jika membeli jenis air minum 1SQ0 maka akan membeli jenis air minum K. 200 ML | 79/124 | 61.7% |
| 3 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum K. 200 ML | 138/219 | 63% |
| 4 | Jika membeli jenis air minum Q FREES maka akan membeli jenis air minum K.200 ML | 36/57 | 63.2% |
| 5 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum D. 220 ML | 140/219 | 63.92% |
| 6 | Jika membeli jenis air minum D.600 ML maka akan membeli jenis air minum D.220 ML | 217/336 | 64.58% |
| ... | | ... | ... |
| 52 | Jika membeli jenis air minum S.200 ML dan GMES, maka akan membeli jenis air minum D.220 ML | 107/128 | 83.59% |
| 53 | Jika membeli jenis air minum D.600 ML dan GMES, maka akan membeli jenis air minum D.220 ML | 41/49 | 83.67% |
| 54 | Jika membeli jenis air minum K.600 ML dan 1SQ0, maka akan membeli jenis air minum K. 200 ML | 27/31 | 87.09% |
| 55 | Jika membeli jenis air minum GMES dan COCO, maka akan membeli jenis air minum D.220 ML | 29/33 | 87.87% |

Hasilnya adalah terbentuk 55 aturan asosiasi yang nilainya lebih atau sama dengan *minimum confidence* yaitu 60 %. Aturan asosiasi dianggap berguna jika nilai

lift ratio ≥ 1 . Semakin tinggi lift ratio, semakin kuat hubungannya. Hasil perhitungan menunjukkan ada 55 kombinasi item yang memenuhi kriteria minimum *confidence* 60% dan lift ratio lebih dari 1, yang berarti aturan tersebut valid.

Tabel 7. Nilai Lift Ratio Aturan Asosiasi

| No | Rule | Confidence (%) | Lift |
|-----|--|----------------|-------|
| 1 | Jika membeli jenis air minum KAS/212 maka akan membeli jenis air minum K. 200 ML | 61.3% | 1.175 |
| 2 | Jika membeli jenis air minum 1SQ0 maka akan membeli jenis air minum K. 200 ML | 61.7% | 1.183 |
| 3 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum K. 200 ML | 63% | 1.208 |
| 4 | Jika membeli jenis air minum Q FREES maka akan membeli jenis air minum K.200 ML | 63.2% | 1.211 |
| 5 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum D. 220 ML | 63.92% | 1.021 |
| 6 | Jika membeli jenis air minum D.600 ML maka akan membeli jenis air minum D.220 ML | 64.58% | 1.031 |
| ... | | ... | ... |
| 52 | Jika membeli jenis air minum S.200 ML dan GMES, maka akan membeli jenis air minum D.220 ML | 83.59% | 1.334 |
| 53 | Jika membeli jenis air minum D.600 ML dan GMES, maka akan membeli jenis air minum D.220 ML | 83.67% | 1.336 |

| No | Rule | Confidence (%) | Lift |
|----|---|----------------|-------|
| 54 | Jika membeli jenis air minum K.600 ML dan 1SQ0, maka akan membeli jenis air minum K. 200 ML | 87.09% | 1.670 |
| 55 | Jika membeli jenis air minum GMES dan COCO, maka akan membeli jenis air minum D.220 ML | 87.87% | 1.403 |

6. Evaluation

Evaluasi Algoritma Apriori dilakukan dengan menetapkan nilai minimum support dan confidence untuk menemukan itemset yang relevan. Proses ini menghasilkan 55 aturan asosiasi dari 5.809 transaksi, yang kemudian disaring berdasarkan nilai support $\geq 3\%$, confidence $\geq 60\%$, dan lift > 1 . Dari hasil tersebut, delapan aturan memenuhi semua kriteria, menunjukkan validitas dan manfaatnya.

Tabel 8. Aturan Asosiasi Final

| No | Rule | Support | Confidence | Lift |
|----|---|---------|------------|-------|
| 1 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum K. 200 ML | 3.2% | 63% | 1.208 |
| 2 | Jika membeli jenis air minum COCO maka akan membeli jenis air minum D. 200 ML | 3.3% | 63.9% | 1.021 |
| 3 | Jika membeli jenis air minum K.200 ML maka akan membeli jenis air minum D. 220 ML | 3.26% | 79.09% | 1.263 |
| 4 | Jika membeli jenis air minum S.200 ML dan K.600 ML maka akan membeli jenis air minum K.200 ML | 4.17% | 70.47% | 1.351 |

| No | Rule | Support | Confidence | Lift |
|----|--|---------|------------|-------|
| 5 | Jika membeli jenis air minum D.220 ML dan K.600 ML maka akan membeli jenis air minum K. 200 ML | 4.8% | 65.2% | 1.250 |
| 6 | Jika membeli jenis air minum D.600 ML maka akan membeli jenis air minum D.220 ML | 5.1% | 64.58 % | 1.031 |
| 7 | Jika membeli jenis air minum GMES maka akan membeli jenis air minum D.220 ML | 5.7% | 78.45 % | 1.252 |
| 8 | Jika membeli jenis air minum K.600 ML maka akan membeli jenis air minum K.200 ML | 10.1 % | 70.6% | 1.354 |

Aturan ini dapat digunakan untuk merancang strategi penjualan yang lebih efektif. Meningkatkan nilai *minimum support* dan *confidence* akan memperkuat asosiasi dan akurasi aturan, memberikan dasar yang lebih kuat untuk strategi bisnis di masa depan.

7. Knowledge

Tahap akhir KDD menampilkan hasil Algoritma Apriori dengan RapidMiner. Proses ini menggunakan data nominal, di mana angka 1 menunjukkan positif dan angka 0 menunjukkan negatif. RapidMiner memproses dan menampilkan hasil itemset beserta nilai *support* dan *confidence* secara otomatis, sesuai dengan perhitungan manual.

Selain *support* dan *confidence*, kekuatan aturan asosiasi dievaluasi menggunakan *lift ratio*. Nilai *lift ratio* > 1 menunjukkan hubungan signifikan antara pembelian item X dan Y. RapidMiner otomatis menghitung nilai *lift ratio* untuk semua pola kombinasi, memudahkan analisis dan visualisasi, seperti yang ditunjukkan pada gambar 5.

| Conclusion | Premises | Support ↑ | Confidence | Lift |
|------------|-------------------|-----------|------------|-------|
| D.220 ML | D.600 ML | 0.017 | 0.783 | 1.250 |
| D.220 ML | NONO | 0.017 | 0.793 | 1.297 |
| K.200 ML | KASI/212 | 0.018 | 0.613 | 1.175 |
| K.200 ML | ISQ0 | 0.018 | 0.617 | 1.183 |
| D.220 ML | KASI/212 | 0.021 | 0.726 | 1.159 |
| D.220 ML | ICE TEA | 0.021 | 0.676 | 1.080 |
| D.220 ML | ISQ0 | 0.024 | 0.789 | 1.260 |
| D.220 ML | S.200 ML_GMES | 0.025 | 0.836 | 1.335 |
| K.200 ML | COCO | 0.032 | 0.630 | 1.208 |
| D.220 ML | COCO | 0.033 | 0.639 | 1.021 |
| D.220 ML | K.200 ML_GMES | 0.033 | 0.791 | 1.283 |
| K.200 ML | S.200 ML_K.600 ML | 0.042 | 0.705 | 1.351 |
| K.200 ML | D.220 ML_K.600 ML | 0.048 | 0.662 | 1.250 |
| D.600 ML | D.600 ML | 0.051 | 0.646 | 1.031 |
| D.220 ML | GMES | 0.057 | 0.785 | 1.253 |
| K.200 ML | K.600 ML | 0.101 | 0.706 | 1.354 |

Gambar 5. Hasil Pola Asosiasi Pada Aplikasi RapidMiner

D. PENUTUP

Berdasarkan penelitian ini, dapat disimpulkan bahwa Algoritma Apriori berhasil mengidentifikasi 55 aturan asosiasi signifikan dalam data penjualan PT Sila Tirta Gemilang. Namun hanya 8 aturan asosiasi yang sudah pasti akan terjadi dikedepannya. Hal ini membantu memahami pola pembelian konsumen dan mendukung pengambilan keputusan bisnis yang lebih tepat.

Identifikasi pola pembelian ini memungkinkan perusahaan merancang strategi penjualan yang lebih efektif, seperti promosi *bundling*. Contohnya, GMES dan COCO yang cenderung dibeli bersama D. 600 ML (*confidence* 87.87%), serta mengoptimalkan manajemen stok untuk mengurangi biaya dan penumpukan barang.

Untuk penelitian selanjutnya, diharapkan dapat menggunakan beberapa algoritma lain sebagai pembanding, seperti algoritma FP-Growth, untuk mengevaluasi hasil yang diperoleh. Selain itu, dapat mencoba berbagai tools selain RapidMiner, seperti WEKA atau aplikasi data mining lainnya, untuk melihat apakah ada perbedaan dalam hasil analisis.

Penelitian selanjutnya juga disarankan untuk menggabungkan data yang mungkin memiliki distribusi yang tidak merata. Penggabungan data dapat membantu mengurangi permasalahan ini, menghasilkan representasi data yang lebih

seimbang dan meningkatkan akurasi hasil analisis.

E. DAFTAR PUSTAKA

- Anas, S., Rumui, N., Roy, A., & Saputro, P. H. (2022). Comparison of Apriori Algorithm and FP-Growth in Managing Store Transaction Data. *International Journal of Computer and Information System (IJCIS)*, 3(4), 158–162. <https://doi.org/10.29040/ijcis.v3i4.96>
- Dora, M., Khairul, R., & Sari, W. M. (2023). Analisa Transaksi Penjualan Dalam peningkatan Promosi Penjualan Berbasis Sistem Informasi. *Ekombis Review: Jurnal Ekonomi Dan Bisnis*, 11(1), 357 – 368. <https://doi.org/10.37676/ekombis.v11i1.2993>
- Erwansyah, K., Andika, B., & Gunawan, R. (2021). Implementasi Data Mining Menggunakan Asosiasi Dengan Algoritma Apriori Untuk Mendapatkan Pola Rekomendasi Belanja Produk Pada Toko Avis Mobile. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 4(1), 148–161. <https://doi.org/10.53513/jsk.v4i1.2628>
- Firdaus, A. A., Iksan, N., Sadih, D. N., Sagita, L., & Setiawan, D. (2021). Penerapan Algoritma Apriori untuk Prediksi Kebutuhan Suku Cadang Mobil. *Justin: Jurnal Sistem Dan Teknologi Informasi*, 9(1), 13–18. <https://doi.org/10.26418/justin.v9i1.41151>
- Gumelar, A. M., & Akbar, M. (2020). Implementasi Data Mining Pengambilan Keputusan Penentuan Korelasi Produk Berdasarkan Pola Transaksi Penjualan Menggunakan Algoritma Apriori. *Prosiding Seminar Multimedia Dan Artificial Intelligence*, 7–14.
- Ikhwan, A., & Aslami, N. (2020). Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means. *Jurnal Teknologi Informasi*, 4(2), 208–217. <https://doi.org/10.36294/jurti.v4i2.2103>
- Kurnia, A., Mirza, A. H., & Andri, A. (2020). Penerapan Decision Tree Data Mining Pada Produksi Kelapa Sawit PT Hindoli Di Sungai Lilin Kabupaten Musi Banyuasin. *Jurnal Pengembangan Sistem Informasi Dan Informatika*, 1(2), 84–99. <https://doi.org/10.47747/jpsii.v1i2.168>
- Riszky, A. R., & Sadikin, M. (2019). Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan. *Jurnal Teknologi Dan Sistem Komputer*, 7(3), 103–108. <https://doi.org/10.14710/jtsiskom.7.3.2019.103-108>
- Syahputri, N. (2020). Penerapan Data Mining Asosiasi pada Pola Transaksi dengan Metode Apriori. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 4(2), 728–736. <https://doi.org/10.30645/j-sakti.v4i2.345>
- Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic: Jurnal Pendidikan Informatika*, 4(1), 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>
- Wulandhari, P. R., Rahaningsih, N., Ali, I., & Rohmat, C. L. (2023). Penerapan Metode Asosiasi Untuk Menemukan Pola Transaksi Penjualan Obat Menggunakan Algoritma FP-Growth. *JATI: Jurnal Mahasiswa Teknik Informatika*, 7(1), 619–630. <https://doi.org/10.36040/jati.v7i1.6404>
-