

PENERAPAN ASSOCIATION RULES MENGGUNAKAN ALGORITMA APRIORI UNTUK IDENTIFIKASI POLA PEMBELIAN

Arya Wijaya¹, Ahmad Faqih², Dodi Solihudin³, Cep Lukman Rohmat⁴, Sandy Eka Permana⁵

^{1,2,3} Program Studi Teknik Informatika, STMIK IKMI, Cirebon, Indonesia

⁴ Program Studi Rekayasa Perangkat Lunak, STMIK IKMI, Cirebon, Indonesia

⁵ Program Studi Manajemen Informatika, STMIK IKMI, Cirebon, Indonesia

Jln. Perjuangan No 10B Majasem, Kesambi Karya Mulya, Kota Cirebon, Jawa Barat Indonesia

aryajt95@gmail.com

ABSTRAK

Transaksi penjualan dalam dunia bisnis seringkali mengalami peningkatan penjualan yang melonjak, dalam hal ini perusahaan-perusahaan memiliki data penjualan bulanan yang terarsip, Masalah yang seringkali terjadi yaitu tidak teratur nya pada penyesuaian stok barang. Data Mining adalah sebuah kegiatan pengolahan data untuk mengenali pola pembelian atau kecenderungan pelanggan terhadap produk penjualan. Penelitian ini bertujuan melibatkan pertanyaan tentang bagaimana algoritma Apriori dapat diterapkan secara efektif untuk mengungkap pola pembelian yang signifikan dalam data transaksi penjualan. studi ini mendiskusikan pentingnya pemahaman pola pembelian dalam bisnis, termasuk manfaatnya untuk rekomendasi produk, strategi pemasaran, dan pengelolaan persediaan. Metode yang digunakan adalah pengumpulan data transaksi penjualan, dengan menggunakan metode KDD dengan algoritma apriori tahapan yang dilakukan antara nya seleksi data, *preprocessing data*, *transformasi data*, *Data mining*, evaluasi. Penggunaan algoritma Apriori untuk mengidentifikasi itemset yang sering muncul bersama dalam transaksi, serta menentukan ambang batas dukungan yang relevan. Hasil dan Diskusi menyoroti pola pembelian yang berhasil diidentifikasi dan menggambarkan implikasinya dalam konteks bisnis, seperti peningkatan penjualan, efisiensi persediaan, dan pengembangan strategi pemasaran yang lebih cermat. Untuk menguji hasil penelitian ini digunakan nilai minimum *Support* 3% dan nilai *Confidence* 50% serta tahap evaluasi menggunakan *lift ratio* minimum 100% dari hasil pengujian tersebut terdapat beberapa yang memenuhi syarat aturan asosiasi. Misalnya kombinasi, *medialuna* dan *coffee* yang memiliki nilai *Support* 3,5%, nilai *Confidence* 57% dan nilai *lift ratio* 120%, Implikasi dari hasil penemuan pola pembelian ini bisa dijadikan strategi penjualan Contohnya, saran untuk penempatan produk, promo khusus untuk produk dan manajemen stok yang efisien

Kata kunci: Asosiasi data, Penjualan barang, Algoritma Apriori, Pola pembelian, Dukungan, Transaksi

1. PENDAHULUAN

Pesatnya perkembangan di bidang Informatika, khususnya dalam hal pengumpulan, penyimpanan, dan analisis data, telah mengubah paradigma dalam berbagai aspek kehidupan. Kemajuan teknologi informasi telah memberikan dampak signifikan pada berbagai sektor, termasuk bisnis, pendidikan, pelayanan kesehatan, dan lain-lain. Data telah menjadi salah satu aset paling berharga dalam era digital ini, dan analisis data telah menjadi komponen penting dalam pengambilan keputusan yang efektif. Di antara berbagai aspek penggunaan data, analisis asosiasi data dalam konteks penjualan barang telah menjadi fokus yang semakin mendalam dalam upaya untuk memahami perilaku konsumen, merencanakan strategi pemasaran yang lebih cermat, dan meningkatkan efisiensi operasional. Dalam kerangka ini, penelitian ini akan mengeksplorasi penerapan algoritma Apriori untuk mengidentifikasi pola pembelian dalam data penjualan barang, dengan harapan memberikan wawasan yang berharga untuk peningkatan kinerja bisnis dan kepuasan pelanggan.

Tiga penelitian *Data Mining* dalam menggali informasi berharga dari data transaksi penjualan dan perilaku konsumen. Dalam kasus 212 Mart, analisis aturan asosiasi dengan algoritma Apriori membantu

memahami pola belanja konsumen dan mengidentifikasi produk-produk yang memiliki korelasi tinggi dalam pembelian bersamaan, memberikan pandangan yang berharga bagi toko dalam mengoptimalkan strategi penjualan. [1] Di Apotek UII Farma, *Data Mining* digunakan untuk menganalisis data penjualan obat, memberikan wawasan yang penting dalam meningkatkan efisiensi dan layanan kepada mahasiswa dan masyarakat. [2] Sementara itu, di Toko Yoss Elektronik, perbandingan antara algoritma Association Rules seperti Apriori dan FP-Growth membantu toko dalam memahami aturan asosiasi yang berbeda, memungkinkan mereka untuk mengambil keputusan yang lebih baik terkait penataan barang dan strategi penjualan. [3] Dalam keseluruhan, ketiga penelitian tersebut menunjukkan betapa pentingnya *Data Mining* dan analisis aturan asosiasi dalam mendukung pengambilan keputusan yang lebih cerdas dan strategi penjualan yang lebih efektif di berbagai jenis bisnis.

Tujuan utama dari penelitian ini adalah menerapkan algoritma Apriori untuk mengidentifikasi pola pembelian dalam data penjualan barang dengan akurat dan efisien. Dengan demikian, penelitian ini bertujuan untuk memberikan wawasan yang lebih dalam mengenai perilaku pelanggan, pola pembelian

produk, serta hubungan antar produk dalam lingkungan bisnis yang semakin terhubung digital. Selain itu, penelitian ini bertujuan untuk mendukung pengambilan keputusan yang lebih cerdas dalam pengembangan strategi pemasaran yang lebih efektif, pengelolaan persediaan yang lebih tepat waktu, dan rekomendasi produk yang lebih personal. Signifikansi penelitian ini terletak pada kontribusinya pada bidang Informatika dan aplikasi praktis dalam bisnis. Dalam domain Informatika, penelitian ini diharapkan dapat mengisi kesenjangan dalam literatur terkait aplikasi algoritma asosiasi dalam analisis data penjualan barang dengan fokus pada peningkatan efisiensi dan akurasi. Sementara dalam konteks bisnis, hasil penelitian ini dapat memberikan manfaat yang signifikan dengan membantu perusahaan meningkatkan strategi pemasaran, merencanakan persediaan, dan meningkatkan kepuasan pelanggan. Dengan begitu, penelitian ini berpotensi menjadi kontribusi berharga dalam mengatasi tantangan dan mengoptimalkan kinerja bisnis di tengah pesatnya perkembangan teknologi informasi dan e-commerce.

Penelitian ini akan mengadopsi metode pendekatan yang terintegrasi, dimulai dengan pengumpulan data penjualan barang yang komprehensif, termasuk informasi transaksi pelanggan, rincian produk, dan sejarah pembelian. Data ini akan diproses dan dinormalisasi untuk mempersiapkannya untuk analisis. Dalam analisis data, kami akan menerapkan algoritma Apriori untuk mengidentifikasi pola pembelian yang signifikan. Selanjutnya, kami akan mengukur kinerja algoritma ini dengan mempertimbangkan akurasi dan efisiensi dalam mengungkap pola-pola tersebut. Pendekatan ini menggabungkan elemen analisis data, pengolahan data, dan implementasi algoritma, menjadikannya metode yang relevan dalam bidang Informatika. Hasil dari penelitian ini akan memberikan wawasan yang lebih dalam tentang bagaimana teknik Informatika dapat digunakan untuk mengidentifikasi pola pembelian dalam konteks data penjualan barang, dan sejauh mana algoritma Apriori dapat efektif mengatasi permasalahan ini. Selain itu, penelitian ini diharapkan memberikan kontribusi berharga dalam pengembangan pengetahuan di bidang ini serta memberikan panduan praktis bagi perusahaan dalam meningkatkan strategi bisnis mereka.

Hasil penelitian ini memiliki implikasi yang signifikan dalam bidang Informatika dan bisnis. Dengan mencapai tujuan penelitian, penelitian ini berpotensi untuk memperkaya pemahaman kita tentang penggunaan algoritma Apriori dalam analisis data penjualan barang. Implikasi akademisnya termasuk kontribusi terhadap literatur tentang algoritma asosiasi, khususnya dalam skala besar dan dalam lingkungan bisnis yang dinamis. Selain itu, hasil penelitian ini memiliki dampak praktis yang substansial. Perusahaan dapat memanfaatkan temuan ini untuk mengoptimalkan strategi pemasaran mereka, meningkatkan rekomendasi produk kepada

pelanggan, dan mengelola persediaan dengan lebih efisien. Hasil ini juga dapat memberikan dorongan bagi penelitian lanjutan dalam pengembangan teknik analisis data yang lebih canggih serta pengaplikasian algoritma asosiasi dalam berbagai konteks bisnis. Dengan demikian, implikasi dari hasil penelitian ini adalah peningkatan pemahaman tentang peran Informatika dalam meningkatkan efisiensi dan efektivitas pengambilan keputusan strategis di era digital.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Dalam konteks ini, teknik, metode, atau algoritma yang digunakan dalam *Data mining* mencakup beragam pendekatan yang dapat diadaptasi sesuai dengan sasaran yang ingin dicapai dan proses KDD (*Knowledge Discovery in Databases*) secara keseluruhan. Pentingnya pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan spesifik analisis serta pada tahapan keseluruhan proses penemuan pengetahuan dari data tersebut. Dengan demikian, *Data mining* bukan hanya tentang mengekstrak pola, tetapi juga melibatkan pemilihan metode yang paling sesuai untuk merinci informasi yang bermanfaat dan relevan sesuai dengan konteks yang diinginkan. [3]

2.2. Aturan Asosiasi

Aturan Asosiasi, yang juga dikenal sebagai *Association rules*, merupakan salah satu teknik dalam data mining yang memiliki kegunaan khusus dalam mengidentifikasi aturan asosiatif terkait dengan suatu set barang. Proses pembentukan aturan asosiasi ini melibatkan analisis pola data yang sering muncul, yang dikenal sebagai *Frequent patterns*, dengan memanfaatkan parameter *Support* dan *Confidence*. *Support* mencerminkan seberapa sering suatu item muncul dalam suatu basis data, sementara *Confidence* menunjukkan seberapa sering pernyataan atau asosiasi tersebut benar. Dengan menggunakan pendekatan ini, aturan asosiasi dapat dihasilkan untuk mengidentifikasi hubungan yang signifikan dan penting antar barang atau elemen dalam suatu konteks data tertentu. [5] *Support* dapat dinyatakan sebagai berikut:

$$Support = \frac{\text{Total Transaksi X dan Y}}{\text{Total Transaksi}} \quad [6]$$

Untuk mencari nilai *Confidence* dapat menggunakan rumus sebagai berikut

$$Confidence = \frac{\text{Total Transaksi X dan Y}}{\text{Total Transaksi X}} \times 100 \quad [6]$$

2.3. Algoritma Apriori

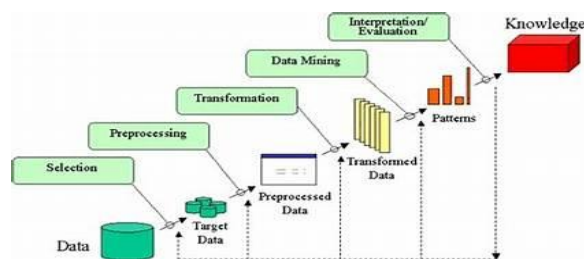
Analisis asosiasi, atau *association rule mining*, adalah suatu teknik dalam bidang *data mining* yang bertujuan untuk mengidentifikasi aturan kombinasi

item. Salah satu tahap penting dalam analisis asosiasi yang mendapat perhatian besar dari peneliti adalah analisis pola frekuensi tinggi (*frequent pattern mining*). *Support* merupakan nilai penunjang atau persentase kemunculan suatu kombinasi item dalam *database*, sementara *confidence* adalah nilai kepastian yang mengindikasikan seberapa kuat hubungan antar item dalam suatu aturan apriori. Setelah pola frekuensi tinggi ditemukan, *confidence* dapat dihitung untuk menentukan kepastian dari hubungan tersebut.

Ketika mengevaluasi keberhasilan suatu asosiasi, dua parameter kritis yang digunakan adalah *support* dan *confidence*. *Support* mencerminkan persentase kemunculan suatu kombinasi item dalam *database*, sedangkan *confidence* menunjukkan seberapa kuat hubungan antar item dalam aturan asosiasi. Dengan mengukur kedua parameter ini, dapat ditentukan sejauh mana suatu asosiasi dianggap signifikan atau tidak. Dengan demikian, metodologi dasar analisis asosiasi melibatkan pencarian pola frekuensi tinggi, pengukuran *support*, dan perhitungan *confidence* untuk mengevaluasi kekuatan hubungan antar item dalam aturan asosiasi yang dihasilkan.

3. METODE PENELITIAN

Metode pengembangan *data mining* yang diterapkan untuk menganalisis data melibatkan langkah-langkah pada proses tahapan *knowledge discovery in databases* (KDD). Proses ini terdiri dari beberapa tahapan, yakni Data, Seleksi data, *Preprocessing*, Transformasi, *Data Mining*, dan Evaluasi. [7]



Gambar 1. Tahapan KDD

Sumber data yang digunakan dalam penelitian ini diperoleh dari Kaggle, sebuah platform terkemuka dalam komunitas *data science* yang menyediakan akses kepada peneliti, ilmuwan data, dan praktisi di berbagai bidang. Data yang digunakan yakni data transaksi penjualan, Dari data yang sudah ada dilanjutkan dengan proses preparasi data melalui tahapan data *selection*, data *cleaning*, dan data *transformation*.

3.1. Seleksi Data

Pemilihan data (seleksi data) adalah proses yang melibatkan penentuan *dataset* atau *subset* dari data yang relevan dan signifikan untuk tujuan analisis atau penelitian tertentu. Pada tahap ini, peneliti atau analis

data memutuskan data mana yang akan digunakan atau diabaikan berdasarkan kriteria tertentu.

3.1.1. Preprocessing Data

Proses pembersihan data (*data cleaning*) adalah tahap di mana dilakukan upaya untuk menghapus *noise* serta data yang tidak konsisten atau tidak relevan. Data yang diperoleh dari basis data suatu perusahaan sering mengandung ketidaksempurnaan, seperti data yang hilang, tidak valid, atau kesalahan pengetikan. Selain itu, terdapat juga atribut-atribut data yang tidak relevan dengan hipotesis *data mining* yang sedang diterapkan. [8]

3.1.2. Transformasi Data

Transformasi Data adalah proses penggabungan data dimana data diubah sesuai dengan kebutuhan *data mining*. [9] Lakukan transformasi data seperti normalisasi atau pengkodean variabel kategoris untuk mempersiapkannya untuk algoritme pemodelan.

3.1.3. Data Mining

Pada penelitian ini, metode pemrosesan data diterapkan melalui teknik *data mining*. *Data mining* merupakan rangkaian langkah untuk menemukan nilai tambah berupa pengetahuan yang sebelumnya tidak diketahui secara manual dari suatu himpunan data. Proses ini memanfaatkan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi serta mengidentifikasi informasi yang berguna dan pengetahuan terkait dari berbagai basis data yang besar. [10]

3.1.4. Evaluasi Model

Evaluasi pola akan dilakukan menggunakan *lift ratio* untuk menilai sejauh mana pentingnya aturan yang dihasilkan oleh program. Nilai *lift ratio* ini akan memberikan indikasi kevalidan informasi terkait apakah pembelian barang X bersamaan dengan barang Y. Informasi ini diperoleh melalui perhitungan *support* dan *confidence* dari aturan yang dihasilkan. Rumus menentukan *lift ratio* dapat menggunakan rumus sebagai berikut:

$$Lift\ Ratio = \frac{Support(X \cap Y)}{Support\ X * Support\ Y} \quad [10]$$

4. HASIL DAN PEMBAHASAN

Analisis data yang digunakan adalah teknik KDD yaitu *Knowledge Discovery in Database*, untuk tahapan dari KDD ini ada beberapa tahapan seperti yang ada dibawah ini:

4.1. Data

Data yang didapatkan berbentuk format excel csv yang bernama bread_basket.csv yang diperoleh dari sumber *dataset* kaggle, data yang berisi berjumlah 20507 transaksi terdiri dari 5 atribut seperti yang ada di gambar 2

Data columns (total 5 columns):

#	Column	Non-Null Count	Dtype
0	Transaction	20507 non-null	int64
1	Item	20507 non-null	object
2	date_time	20507 non-null	object
3	period_day	20507 non-null	object
4	weekday_weekend	20507 non-null	object

Gambar 2. Data dan Atribut

Agar *dataset* bisa diakses di *software google collaborator*, maka perlu dilakukan import data bread basket.csv dengan perintah `df = pd.read_csv("/content/drive/MyDrive/DATASET/bread basket.csv")` seperti yang ada pada gambar 3

```
# mengimport dataset yang akan kita gunakan
df = pd.read_csv("/content/drive/MyDrive/DATASET/bread basket.csv")
df
```

Transaction	Item	date_time	period_day	weekday_weekend
0	1 Bread	30-10-2022 09:58	morning	weekend
1	2 Scandinavian	30-10-2022 10:05	morning	weekend
2	2 Scandinavian	30-10-2022 10:05	morning	weekend
3	3 Hot chocolate	30-10-2022 10:07	morning	weekend
4	3 Jam	30-10-2022 10:07	morning	weekend
...	...	...	...	...
20502	9692 Coffee	9/4/2023 14:32	afternoon	weekend
20503	9682 Tea	9/4/2023 14:32	afternoon	weekend
20504	9683 Coffee	9/4/2023 14:37	afternoon	weekend
20505	9683 Pastry	9/4/2023 14:37	afternoon	weekend
20506	9684 Smoothies	9/4/2023 15:04	afternoon	weekend

Gambar 3. Import Data

4.2. Selection

Setelah data diimport dilakukan seleksi data dimana data yang diseleksi adalah data yang digunakan untuk pengolahan *data mining* sesuai dengan analisis data, proses yang dilakukan meliputi penghapusan atribut data yang tidak digunakan dalam proses analisis data mining dengan perintah *python* seperti digambar 4

```
# menghapus kolom date_time, period_day, weekday_weekend
df = df.drop(columns=['date_time', 'period_day', 'weekday_weekend'])
df
```

Transaction	Item
0	1 Bread
1	2 Scandinavian
2	2 Scandinavian
3	3 Hot chocolate
4	3 Jam
...	...
20502	9692 Coffee
20503	9682 Tea
20504	9683 Coffee
20505	9683 Pastry
20506	9684 Smoothies

Gambar 4. Seleksi Data

4.3. Preprocessing

Proses cleaning data atau pembersihan data yang missing atau memiliki nilai yang tidak konsisten, pada proses ini dilakukan analisa terlebih dahulu apakah atribut data memiliki nilai *missing*, Tujuan dari *preprocessing* adalah untuk membersihkan, mengubah, dan mempersiapkan data mentah agar sesuai untuk analisis lebih lanjut. Adapun cara untuk membersihkan data yang memiliki nilai *missing* dengan menggunakan *software python* atau perintah *python* yang diterapkan dalam *software* tersebut, berikut contoh perintah dari *python* untuk *preprocessing* data

```
[8] # membersihkan kolom item
df['Item'] = df['Item'].str.strip()
df['Item'] = df['Item'].str.lower()

# menampilkan kolom yang sudah dibersihkan
df.head()
```

Transaction	Item
0	1 bread
1	2 scandinavian
2	2 scandinavian
3	3 hot chocolate
4	3 jam

Gambar 5. Preprocessing Data

Setelah data dilakukan pembersihan selanjutnya data perlu ditransformasikan agar terbaca oleh *software* untuk pengolahan *data mining* yang akan dilakukan menggunakan aturan asosiasi rules dengan algoritma *apriori* untuk menemukan pola pembelian pelanggan

4.4. Transformasi Data

Dari data yang diperoleh setelah melewati proses seleksi data dan *preprocessing* data selanjutnya data ditransformasikan sesuai dengan analisis data, Transformasi ini dapat melibatkan pengubahan skala, pengkodean, normalisasi, atau perubahan struktur data agar lebih sesuai untuk analisis atau pemodelan. Data akan diubah label *encoding* yaitu Menggantikan nilai-nilai kategori dengan nilai-nilai *numerik*. ketika label *encoding* menunjukkan angka 0 maka tidak ada transaksi didalam namun label *encoding* menunjukkan nilai 1 maka ada transaksi didalamnya.

Sebelum data diubah menjadi data label *encoding* terlebih dahulu dilakukan penjumlahan data transaksi penjualan dalam kombinasi atribut transaction dan item menggunakan perintah *python*.

```
# menghitung jumlah item untuk setiap kombinasi dari kolom 'Transaction' dan 'Item' menggunakan
transaction_item = df.groupby(['Transaction', 'Item']).count().reset_index()
transaction_item
```

Transaction	Item	Count
0	1 bread	1
1	2 scandinavian	2
2	2 scandinavian	1
3	3 hot chocolate	1
4	3 jam	1
...	...	...
20502	9692 coffee	1
20503	9682 tea	1
20504	9683 coffee	1
20505	9683 pastry	1
20506	9684 smoothies	1

Gambar 6. Membuat table baru count

Adapun untuk perhitungan menu lainnya dilakukan dengan cara yang sama dan diperoleh informasi *confidence* menggunakan perintah *python*. Setelah itu baru data diubah menjadi label *encoding* menggunakan perintah *python* seperti digambar 7

```
# membuat matrix non binary transaksi dan item dan setiap baris mewakili apakah item tersebut ada dalam transaksi atau tidak
xy_basket = transaction_item.pivot(index='Transaction', columns='Item', values='count').fillna(0)
xy_basket.head()
```

Transaction	adjustment	afternoon with the baker	afternoon	afternoon right at ray	bacon	baguette	bakewell	bare popcorn	basket	beefsteak pie	bread	bread pudding
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Gambar 7. Transformasi Data

4.5. Data Mining

Data mining adalah bagian dari proses *Knowledge Discovery in Database (KDD)*, yang melibatkan analisis data dalam jumlah besar. Proses data mining menggunakan berbagai metode untuk

menggali informasi dari data dan menemukan model-model baru yang tersembunyi di dalamnya. Model-model baru ini berguna untuk memahami informasi yang terkandung dalam dataset yang besar tersebut. Penerapan data mining dapat diterapkan dalam berbagai disiplin ilmu yang melibatkan pengolahan data, seperti statistika, kesehatan, basis data, dan bidang lainnya.

Peneliti menggunakan metode algoritma data mining, khususnya Apriori, untuk mengolah data dalam rangka mendapatkan hasil yang dapat dijadikan rekomendasi dalam konteks bisnis penjualan. Pemilihan algoritma Apriori didasarkan pada kemampuannya untuk menentukan produk yang sering terjual dan memberikan rekomendasi untuk produk-produk yang saling terkait dan mungkin akan dibeli oleh pelanggan. Proses analisis asosiasi terdiri dari dua tahap utama, yaitu a) analisis pola frekuensi tertinggi dengan mencari kombinasi item yang memenuhi syarat minimum skor *support*, dan b) pembentukan aturan asosiasi yang memenuhi syarat minimum untuk *confidence*.

4.5.1. Analisis Frekuensi Tertinggi

Tahapan ini merupakan langkah awal dalam pengelolaan data transaksi menggunakan Algoritma Apriori. Pada tahapan ini, terdapat parameter khusus yang digunakan untuk menentukan kriteria batasan analisis, yaitu nilai minimum *support*. Rentang nilai minimum *support* berkisar antara 0 hingga 1, dan dalam penelitian ini, nilai minimum *support* telah ditetapkan sebesar 0,03 atau 3%. Penentuan nilai minimum *support* dilakukan agar fokus analisis dapat diarahkan kepada pola yang memiliki dampak signifikan. Nilai *support* yang memenuhi batas minimum digabungkan dengan *itemset* lain, dan nilai *support* dihitung ulang hingga tidak ada lagi kombinasi *itemset* yang dapat terbentuk. *Itemset* yang tidak memenuhi nilai minimum *support* akan dieliminasi dan tidak digunakan pada tahapan selanjutnya untuk mencari nilai *confidence* dan nilai lift ratio. Berikut adalah hasil nilai *support* untuk kombinasi 1-*Itemset* menggunakan perintah *python* pada gambar 8

```
# artinya item tersebut harus ada dalam minimal 282 transaksi dari 9465 transaksi jika
# frequent itemset
frequent_itemsets = apriori(my_basket_sets, min_support = 0.03, use_colnames = True)

# Menampilkan frequent itemsets
print("Frequent 1-itemset:")
print(frequent_itemsets[frequent_itemsets['itemsets'].apply(len) == 1])
```

	support	itemsets
0	0.036344	(alfajores)
1	0.327205	(bread)
2	0.040042	(brownie)
3	0.103856	(cake)
4	0.478394	(coffee)
5	0.054411	(cookies)
6	0.039197	(farm house)
7	0.058320	(hot chocolate)
8	0.038563	(juice)
9	0.061807	(medialuna)
10	0.038457	(muffin)
11	0.086107	(pastry)
12	0.071844	(sandwich)
13	0.034548	(scone)
14	0.034443	(soup)
15	0.142631	(tea)
16	0.033597	(toast)

Gambar 8. frequent 1-Itemset

Selanjutnya dari hasil perintah *python* kita buat hasil mencari *frequent 1-Itemset* menjadi tabel sesuai dengan hasil yang didapatkan pada Tabel 1

Tabel 1. Support 1-itemset

Transaction	Itemsets	Support
0	(alfajores)	0.036344
1	(bread)	0.327205
2	(brownie)	0.040042
3	(cake)	0.103856
4	(coffee)	0.478394
5	(cookies)	0.054411
6	(farm house)	0.039197
7	(hot chocolate)	0.058320
8	(juice)	0.038563
9	(medialuna)	0.061807
10	(muffin)	0.038457
11	(pastry)	0.086107
12	(sandwich)	0.071844
13	(scone)	0.034548
14	(soup)	0.034443
15	(tea)	0.142631
16	(toast)	0.033597

Tabel 1 menunjukkan pola kombinasi 1-*Itemset* yang memenuhi nilai minimum *support*. Contohnya adalah "coffee" dengan nilai *support* sebesar 0.478394, yang menjadi pola frekuensi tinggi dalam pembentukan kombinasi 1-*Itemset*. Proses perhitungan untuk menu lain dilakukan dengan metode yang serupa, dan diperoleh informasi *support* untuk 2-*Itemset* menggunakan persamaan 2.

Sementara itu, perhitungan untuk menu lainnya juga dilakukan dengan pendekatan yang sama, dan hasilnya memberikan informasi *support* untuk 2-*Itemset* menggunakan perintah *python* pada gambar 4.8.

```
# menggunakan 'algoritma apriori' dengan min_support=0.03 (3% dari 9465)
# artinya item tersebut harus ada dalam minimal 282 transaksi dari 9465 transaksi jika tidak memenuhi
# frequent itemset
frequent_itemsets = apriori(my_basket_sets, min_support = 0.03, use_colnames = True)

# Menampilkan frequent itemsets
print("Frequent 2-itemset:")
print(frequent_itemsets[frequent_itemsets['itemsets'].apply(len) == 2])
```

	support	itemsets
17	0.090016	(bread, coffee)
18	0.054728	(cake, coffee)
19	0.035182	(medialuna, coffee)
20	0.047544	(pastry, coffee)
21	0.038246	(sandwich, coffee)
22	0.049868	(coffee, tea)

Gambar 9. Frequent 2-Itemset

Selanjutnya dari hasil perintah *python* kita buat hasil mencari *frequent 2-Itemset* menjadi tabel sesuai dengan hasil yang didapatkan pada Tabel 2.

Tabel 2. Support 2-Itemset

Transaction	Itemsets	Support
0	(coffee, bread)	0.090016
1	(cake, coffee)	0.054728
2	(coffee, medialuna)	0.035182
3	(pastry, coffee)	0.047544
4	(sandwich, coffee)	0.038246
5	(tea, coffee)	0.049868

Tabel 2 menampilkan enam pola kombinasi 2-Itemset yang memenuhi nilai minimum *support*. Salah satu contohnya adalah kombinasi (coffee, bread) dengan nilai *support* sebesar 0,090016, yang menjadi pola frekuensi tinggi dalam pembentukan Kombinasi 2-Itemset. Proses perhitungan untuk kombinasi menu lainnya dilakukan dengan metode yang serupa, dan diperoleh informasi *support* untuk 3-Itemset menggunakan persamaan 3 menggunakan perintah *python* pada gambar 10.

```

# menggunakan 'algoritma apriori' dengan min_support=0.03 (3% dari 9465)
# artinya item tersebut harus ada dalam minimal 282 transaksi dari 9465 transaksi jika tidak
# frequent itemset
frequent_itemsets = apriori(my_basket_sets, min_support = 0.03, use_colnames = True)

# Menampilkan frequent itemsets
print("Frequent 3-Itemset:")
print(frequent_itemsets[frequent_itemsets['itemsets'].apply(len) == 3])

Frequent 3-Itemset:
Empty DataFrame
Columns: [support, itemsets]
Index: []
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `shou

```

Gambar 10. Frequent 3-Itemset

Selanjutnya dari hasil perintah *python* kita buat hasil mencari *frequent 3-Itemset* menjadi tabel sesuai dengan hasil yang didapatkan pada Tabel 3

Tabel 3. Support 3-Itemset

Transaction	Itemsets	Support
0	Tidak ada data yang memenuhi min support	Tidak ada kombinasi yang memenuhi nilai min support

Tabel 3 tidak menunjukkan adanya pola kombinasi 3-Itemset karena tidak ada yang memenuhi nilai minimum *support*. Oleh karena itu, hanya pola kombinasi 2-Itemset yang memenuhi kriteria nilai minimum *support* untuk membentuk aturan asosiasi.

4.5.2. Pembentukan Aturan Asosiasi

Setelah menemukan semua pola frekuensi tinggi, langkah berikutnya adalah mencari aturan asosiasi yang memenuhi persyaratan nilai minimum *confidence*, yang telah ditetapkan pada 0,5 atau 50% dari kombinasi 2-Itemset yang telah diidentifikasi sebelumnya. Pemilihan nilai minimum *confidence* dilakukan untuk memastikan bahwa aturan asosiasi yang terbentuk memiliki tingkat kepercayaan yang memadai. Jika nilai *confidence* yang dihasilkan tidak mencapai nilai minimum yang ditetapkan, aturan tersebut akan dihilangkan atau dieliminasi. Dalam beberapa kasus, mungkin terjadi perubahan posisi item pada beberapa kombinasi, yang disebabkan oleh fakta bahwa aturan asosiasi yang salah satunya lebih umum dalam *dataset* umumnya akan memiliki nilai *support* yang lebih tinggi. Adapun cara untuk menentukan nilai minimum *confidence* menggunakan perintah *python* pada gambar 4

```

[19] # sekarang membuat aturan dari frequent itemset yang dihasilkan di atas

rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=0.5)
rules = rules[['antecedents', 'consequents', 'support', 'confidence']]
rules

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning:
and should_run_async(code)

   antecedents  consequents  support  confidence
0      (cake)      (coffee)  0.054728  0.526958
1  (medialuna)      (coffee)  0.035182  0.569231
2      (pastry)      (coffee)  0.047544  0.552147
3  (sandwich)      (coffee)  0.038246  0.532353

```

Gambar 11. Mencari nilai min confidence

Selanjutnya dari hasil perintah *python* kita buat hasil mencari minimum *confidence* pada *frequent 2-Itemset* yang sudah dicari nilai *support* nya menjadi tabel sesuai dengan hasil yang didapatkan pada Tabel 4.

Tabel 4. Confidence

No	Item 1	Item 2	Confidence	Persentase
1	(medialuna)	(coffee)	0.569231	57%
2	(pastry)	(coffee)	0.552147	55%
3	(sandwich)	(coffee)	0.532353	53%
4	(cake)	(coffee)	0.526958	53%

Tabel 4 menunjukkan lima aturan asosiasi yang memenuhi nilai minimum *confidence*, salah satunya adalah (medialuna, Coffee) dengan nilai *confidence* sebesar 0.569231, menjadi aturan asosiasi dengan tingkat kepercayaan tertinggi. Langkah terakhir adalah mencari nilai *lift ratio*, di mana pengujian menggunakan *lift ratio* dilakukan untuk memperkuat kekuatan aturan asosiasi tersebut. Jika nilai *lift ratio* sama dengan atau lebih besar dari 1, hal itu menandakan bahwa aturan asosiasi tersebut kuat. Sebaliknya, jika nilai *lift ratio* kurang dari 1, menunjukkan adanya ketergantungan antara dua itemset atau antara item. Dalam konteks analisis asosiasi, ini mengindikasikan bahwa kedua item cenderung muncul bersamaan lebih jarang. Penting untuk dicatat bahwa interpretasi nilai *lift ratio* tidak dapat dipahami secara terpisah dan harus dilihat bersama-sama dengan nilai *support* dan nilai *confidence*. Berikut hasil nilai kombinasi itemset yang memenuhi syarat min *lift ratio* menggunakan perintah *python* pada gambar 12.

```

rules = association_rules(frequent_itemsets, metric="lift", min_threshold=1)
rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=0.5)
rules = rules[['antecedents', 'consequents', 'support', 'confidence', 'lift']]
rules

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning:
and should_run_async(code)

   antecedents  consequents  support  confidence  lift
0      (cake)      (coffee)  0.054728  0.526958  1.101515
1  (medialuna)      (coffee)  0.035182  0.569231  1.189878
2      (pastry)      (coffee)  0.047544  0.552147  1.154168
3  (sandwich)      (coffee)  0.038246  0.532353  1.112792

```

Gambar 12. Mencari min lift ratio

Selanjutnya dari hasil perintah *python* kita buat hasil mencari minimum *lift ratio* pada *frequent 2-Itemset* yang sudah memenuhi syarat nilai *support* dan *confidence* nya menjadi tabel sesuai dengan hasil yang didapatkan pada Tabel 5

Tabel 5. Lift Ratio

No	Item 1	Item 2	Support	Confidence	Lift
1	(medialuna)	(coffee)	0.035182	0.569231	1.189878
2	(pastry)	(coffee)	0.047544	0.552147	1.154168
3	(sandwich)	(coffee)	0.038246	0.532353	1.112792
4	(cake)	(coffee)	0.054728	0.526958	1.101515

4.6. Evaluasi

Dari hasil pengelolaan dan pengujian data transaksi, salah satu contoh kombinasi *itemset* yang memiliki nilai *support* dan nilai *confidence* adalah (medialuna, Coffee). Kombinasi ini memiliki nilai *support* sebesar 0,035182 atau 3,5%, yang berarti bahwa 3,5% dari seluruh transaksi mengandung kedua item tersebut. Selain itu, nilai *confidence*-nya sebesar 0,569231 atau 57%, menunjukkan bahwa dari transaksi yang mengandung kedua item, sebanyak 57% juga mengandung medialuna. Tabel 4.6 memperlihatkan hasil penarikan kesimpulan, yang mencakup kombinasi 2 *itemset* beserta nilai *support*, *confidence*, dan *lift ratio* setiap kombinasi menu.

Tabel 6. Evaluasi

Item 1	Item 2	Algoritma Apriori		
		Lift	Support	Confidence
(medialuna)	(coffee)	1.189878	0.035182	0.569231
(pastry)	(coffee)	1.154168	0.047544	0.552147
(sandwich)	(coffee)	1.112792	0.038246	0.532353
(cake)	(coffee)	1.101515	0.054728	0.526958

5. Kesimpulan dan Saran

Berdasarkan hasil analisis menggunakan algoritma Apriori pada data transaksi, ditemukan pola pembelian yang signifikan antara dua produk, yaitu medialuna dan kopi. Kombinasi kedua item ini muncul dalam 3,5% dari total transaksi, menunjukkan bahwa pelanggan memiliki kecenderungan untuk membeli kedua produk tersebut secara bersamaan. Dengan nilai *confidence* sebesar 57%, dapat disimpulkan bahwa kemungkinan besar, jika pelanggan membeli medialuna, mereka juga akan membeli kopi dalam satu transaksi. Hasil ini memberikan wawasan yang berharga terkait preferensi pelanggan dan interaksi antara produk, yang dapat dimanfaatkan untuk meningkatkan strategi penjualan dan pemasaran. Penggunaan parameter aturan asosiasi, seperti nilai minimum *support* sebesar 3%, nilai *confidence* sebesar 50%, dan *lift ratio* 1, memberikan pedoman dalam mengidentifikasi pola pembelian yang relevan. Informasi ini dapat digunakan oleh bisnis untuk mengoptimalkan penempatan produk, merancang promosi khusus, dan meningkatkan kepuasan pelanggan, sehingga mendukung pengambilan keputusan yang lebih cerdas dalam ranah penjualan dan pemasaran. Dari evaluasi penerapan algoritma Apriori, dapat disimpulkan bahwa analisis *matriks support*, *confidence*, dan *lift ratio* membantu mengidentifikasi aturan asosiasi yang kuat dalam data transaksi. Nilai-nilai *support* dan *confidence* yang tinggi mempersempit jumlah aturan yang dihasilkan,

menunjukkan relevansi dan keakuratan aturan tersebut. Pencarian kombinasi *itemset* dengan nilai minimum *support* 0,03 menghasilkan 4 aturan asosiasi dengan minimum *confidence* 0,5. Evaluasi menggunakan *lift ratio* mengonfirmasi keterkaitan kuat antar aturan, dengan nilai *lift ratio* melebihi 1. Kesimpulannya, algoritma Apriori efektif dalam mengidentifikasi pola pembelian yang signifikan, memberikan landasan penting untuk merancang strategi penjualan yang lebih baik dan memahami preferensi pelanggan secara lebih mendalam.

Saran untuk peningkatan strategi penjualan melibatkan implementasi *bundling* produk dan penyesuaian penempatan di toko. Disarankan untuk terus memonitor tren pembelian pelanggan dan menyesuaikan paket promosi secara berkala untuk menjaga daya tariknya. Selain itu, menjaga ketersediaan produk yang berkolerasi kuat di dekat satu sama lain di toko dapat meningkatkan kemungkinan *cross-selling*. Dengan mempertahankan responsif terhadap perubahan preferensi pelanggan, bisnis dapat terus memperkuat strategi penjualan dan meningkatkan kepuasan pelanggan.

Studi kasus pada cabang atau outlet tertentu dapat memberikan pemahaman yang lebih spesifik, sementara evaluasi dampak dari strategi penjualan yang diusulkan menjadi fokus utama penelitian. Kolaborasi lebih lanjut dengan pemilik toko dapat memberikan perspektif praktis yang berguna. Dengan pendekatan ini, penelitian berikutnya diharapkan dapat memberikan kontribusi yang lebih komprehensif terhadap pemahaman pola pembelian pelanggan dan pengembangan strategi penjualan yang lebih efektif.

DAFTAR PUSTAKA

- [1] A. Novianti and E. Elisa, "Penentuan Aturan Asosiasi Pola Pembelian Pada Minimarket Dengan Algoritma Apriori," *Build. Informatics, Technol. Sci.*, 2020.
- [2] R. Febrian, F. Dzulfaqor, and M. N. Lestari, "Analisis pola pembelian obat di apotek uii farma menggunakan metode algoritma apriori," *Semnasteknomedia*, 2018, [Online]. Available: <http://ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/2056>
- [3] T. Prasetya, J. E. Yanti, A. I. Purnamasari, A. R. Dikananda, and O. Nurdiawan, "Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori," *INFORMATICS Educ. Prof. J. Informatics*, vol. 6, no. 1, p. 43, 2022, doi: 10.51211/itbi.v6i1.1688.
- [4] D. Maulana and M. Kiptiyah, "Analisa Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Galeri Elzatta Cikarang," *J. Teknol. Pelita Bangsa*, vol. 10, pp. 18–26, 2019.
- [5] I. Qoniah and A. T. Priandika, "Analisis Market Basket Untuk Menentukan Asosiasi Rule

- Dengan Algoritma Apriori (Studi Kasus: Tb. Menara),” *J. Teknol. Dan Sist. Inf.*, 2020, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/sisteminformasi/article/view/368>
- [6] V. C. Nisa and F. N. Khasanah, “Algoritma Apriori Dalam Identifikasi Pola Pembelian Konsumen Pada Produk Minuman,” *INFORMAL Informatics J.*, 2023, [Online]. Available: <https://jurnal.unej.ac.id/index.php/INFORMAL/article/view/41514>
- [7] A. Musaddad, O. Nurdiawan, and G. Dwilestari, “Penerapan Association Rule Menggunakan Frequent Pattern Growth Untuk Rekomendasi Produk Jersey Sepakbola,” *J. Comput. Syst. informatics*, 2022, [Online]. Available: <http://ejurnal.seminar-id.com/index.php/josyc/article/view/1390>
- [8] Y. Apridonal M, W. Choiriah, and A. Akmal, “Penerapan Data Mining Menggunakan Metode Association Rule Dengan Algoritma Apriori Untuk Analisa Pola Penjualan Barang,” *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 5, no. 2, pp. 193–198, 2019, doi: 10.33330/jurteksi.v5i2.362.
- [9] D. Maulana and M. Kiptiyah, “Analisa Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Galeri Elzatta Cikarang,” *J. SIGMA*, 2019, [Online]. Available: <https://www.jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/539>
- [10] N. Hadinata, “Analisis Pola Pembelian Produk Makanan Ringan Menggunakan Algoritma Apriori,” *J. SISFOKOM (Sistem Inf. dan Komputer)*, 2020, [Online]. Available: <http://jurnal.atmaluhur.ac.id/index.php/sisfokom/article/view/623>