

ANALISIS POLA PENJUALAN PRODUK DI E-COMMERCE MEGGUNAKAN ALGORITMA APRIORI (STUDI KASUS: TOKO DFS14)

Nur Safitri, Mohamad Jajuli, Sofi Defiyanti, Garno

Informatika, Universitas Singaperbangsa Karawang

Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia

2010631170025@student.unsika.ac.id

ABSTRAK

DFS14 adalah toko yang berfokus dalam menyediakan kemasan makanan seperti box, kantong kertas, dan berbagai jenis kemasan lainnya yang menggunakan bahan *food grade*. Toko ini beroperasi melalui *platform e-commerce* seperti Shopee untuk transaksi penjualannya. Selama 8 tahun terakhir, toko ini telah mencatat lebih dari 1000 transaksi. Namun, DFS14 sering menghadapi masalah dalam hal penyetokan ulang produk yang tidak sesuai dengan kebutuhan customer yang dapat menyebabkan pembatalan otomatis di *platform e-commerce* dan berdampak pada rating toko. Dengan menganalisis kombinasi produk yang sering dibeli bersamaan, toko dapat melakukan penyetokan ulang sesuai kebutuhan, sehingga mengurangi risiko kehabisan stok. Penelitian ini menggunakan algoritma apriori dan metodologi KDD (*Knowledge Discovery in Database*) untuk menganalisis data transaksi dan menemukan kombinasi itemset pada penjualan produk kemasan. Dari hasil analisis, ditemukan 184 transaksi yang menghasilkan 11 aturan terbaik dengan nilai minimum *support* 40% dan nilai minimum *confidence* 80%. Semua aturan tersebut telah diuji menggunakan evaluasi *lift ratio* dan memenuhi nilai evaluasi yang valid. Hasil terbaik menunjukkan nilai evaluasi *lift ratio* sebesar 1,9 dengan nilai *confidence* sebesar 94,382%. Hasil penelitian ini diharapkan dapat membantu DFS14 dalam mengelola strategi penjualan dan penyetokan produk.

Kata kunci : Algoritma Apriori, Knowledge Discovery in Database, E-commerce

1. PENDAHULUAN

Kemajuan teknologi informasi dan internet memberikan dampak yang signifikan terhadap dunia bisnis. Pertumbuhan *e-commerce* juga didorong oleh perubahan perilaku konsumen. Berdasarkan laporan *We Are Social*, Indonesia berada di urutan keenam dengan 62,6%, masyarakat berbelanja online setiap pekan [1].

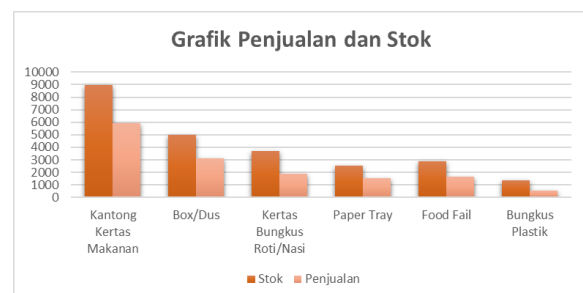
Bagi para pelaku bisnis, *e-commerce* memberikan kesempatan untuk mengakses pasar yang lebih luas tanpa memerlukan investasi besar dalam pembangunan toko fisik [2].

Namun penggunaan *e-commerce* juga dapat mengalami kerugian jika penjualannya tidak tepat sasaran. Ketika stok barang melebihi permintaan, toko akan mengalami kerugian terutama pada barang yang bisa kadaluwarsa atau rusak. Sebaliknya jika stok kurang dari permintaan, konsumen mencari toko lain dan toko kehilangan kesempatan penjualan.

DFS14 merupakan toko yang berfokus pada penyediaan kemasan makanan seperti box, kantong kertas, dan berbagai jenis kemasan lainnya yang terbuat dari bahan *food grade* dan telah memulai kegiatan usahanya sejak tahun 2016. Salah satu *platform e-commerce* yang digunakan yaitu Shopee. Banyaknya data transaksi yang telah dilakukan sering kali hanya digunakan sebagai laporan penjualan saja, padahal dari data transaksi tersebut dapat memberikan informasi yang dapat kita manfaatkan untuk mengetahui hubungan antar item [3].

DFS14 seringkali mengalami kendala pada penyetokan ulang produk dan terkendala dalam melakukan strategi penjualan, salah satu contohnya ketika customer ingin membeli kantong fc dan kertas nasi cetak namun yang tersedia hanya kantong fc dengan box donat.

Kekurangan atau kekosongan stok di *seller center* pada *e-commerce* dapat menyebabkan pembatalan otomatis pada *platform e-commerce* nya, yang berdampak finalty pada toko dan mengurangi tingkat rating. Hal ini menekankan pentingnya untuk memastikan proses transaksi dilakukan dengan cepat dan pengiriman dilakukan sebelum batas pembatalan otomatis oleh sistem untuk menghindari dampak negatif yang signifikan pada keberhasilan dan kelancaran transaksi jual beli. Grafik di bawah ini menampilkan penjualan dan stok beberapa kategori produk yang paling diminati oleh konsumen (*best seller*). Data yang ditampilkan adalah hasil perhitungan produk selama 12 bulan.



Gambar 1. Grafik penjualan dan stok tahun 2023

Grafik pada gambar diatas Menunjukkan adanya selisih antara penjualan dan persediaan stok yang menyebabkan penumpukan produk. Untuk mengatasi masalah diatas, diperlukan analisis penjualan lebih lanjut untuk mengetahui produk yang sering terjual bersamaan. Dengan menganalisis kombinasi produk yang sering dibeli bersamaan, toko dapat melakukan penyetokan ulang sesuai kebutuhan, sehingga mengurangi risiko kehabisan stok. Analisis data penting untuk mendukung pengambilan keputusan dengan sistem data yang tersedia [4].

Data mining, atau Knowledge Discovery in Database (KDD), mengolah data transaksi penjualan untuk menemukan informasi berbasis pengetahuan, mengidentifikasi pola atau relasi dalam data besar. Dalam data mining, metode yang digunakan harus disesuaikan dengan sifat data dan tujuan yang diharapkan. Penelitian oleh Amalia & Setiawansyah pada penjualan handphone menggunakan Algoritma Apriori berhasil menemukan 15 frequent itemset dan aturan asosiasi dengan minimum *support* dan *confidence* yang memenuhi persyaratan [5].

Pada penelitian ini berfokus pada penggunaan algoritma Apriori untuk menganalisis pola penjualan produk di *e-commerce*. Dengan mengkaji nilai *support* dan *confidence*, algoritma ini menghasilkan aturan asosiasi yang membantu memahami pola transaksi produk dan mengembangkan strategi penjualan pada DLS14.

2. TINJAUAN PUSTAKA

2.1. Penjualan

Penjualan adalah upaya terencana untuk memenuhi kebutuhan dan keinginan pembeli guna memperoleh pendapatan. Ini penting bagi pengusaha karena penjualan menghasilkan keuntungan dan membantu mengukur daya tarik produk serta tanggapan pembeli [6].

Dengan demikian, penjualan dianggap sebagai jantung bisnis, dari mana pendapatan diperoleh dan keberhasilan produk di pasar diukur.

2.2. Strategi Pemasaran

Strategi pemasaran melibatkan perencanaan dan pelaksanaan langkah-langkah untuk memahami pasar, mengembangkan produk, menetapkan harga, memilih saluran distribusi, dan merancang promosi guna mencapai tujuan perusahaan. Fokusnya pada pasar, segmen konsumen, nilai produk, dan komunikasi merek. Menurut Marwan, pemasaran online adalah strategi untuk memenuhi kebutuhan konsumen dan menghasilkan keuntungan, dengan penekanan lebih besar pada kepuasan pelanggan, karena konsumen hanya dapat melihat produk melalui gambar[7].

2.3. Strategi Penjualan

Strategi penjualan meliputi perencanaan dan taktik yang dilakukan untuk mencapai target penjualan, memperluas jangkauan pasar, mengelola proses penjualan, dan meningkatkan konversi

pelanggan potensial menjadi pelanggan yang sebenarnya. Lebih terfokus pada aktivitas-aktivitas yang berhubungan langsung dengan proses penjualan, negosiasi, dan pelayanan pelanggan.

2.4. Produk Kemasan

Kemasan produk adalah suatu jenis wadah yang digunakan untuk melindungi, mengemas, mengirimkan, menyimpan, mengidentifikasi, dan membedakan barang-barang yang ada di pasar. Sebagaimana dikutip dalam Firmansyah, 2022, hlm. 176 sebelumnya, produk kemasan hanya digunakan untuk melindungi barang atau memudahkan pengangkutan. Dengan meningkatnya nilai fungsionalnya, kemasan telah diakui sebagai faktor penting dalam persaingan pasar [8].

2.5. E-commerce

Perdagangan dianggap sebagai contoh perubahan gaya hidup masyarakat yang disebabkan oleh kemajuan IT [9].

Diharapkan bahwa e-commerce dapat memberikan manfaat yang penting dalam menghadapi persaingan bisnis yang sengit, dan dunia usaha harus terus mengintegrasikan teknologi ke dalam operasional bisnisnya. *E-commerce* merupakan proses dimana barang, jasa, dan informasi dibeli, dijual, dan ditukar melalui jaringan komputer [2].

2.6. Data Mining

Data mining ialah proses otomatis dengan menggunakan lebih dari satu algoritma *machine learning* untuk menganalisis dan mengekstrak pengetahuan dari kumpulan data [4].

Data mining adalah proses yang melibatkan Penggunaan algoritma dan teknik dari bidang statistik, pembelajaran mesin, dan sistem manajemen basis data terkait dengan *data mining*. Metode dalam data mining, yang juga dikenal sebagai *Knowledge Discovery in Database*, merupakan bidang ilmu yang berfokus pada pengidentifikasian pola dalam data. Data mining merupakan istilah yang digunakan untuk menggambarkan proses-proses yang dilakukan untuk mendapatkan pengetahuan atau pola dari kumpulan data.

2.7. Algoritma Apriori

Algoritma Apriori, yang dikembangkan oleh Agrawal dan Srikant pada tahun 1994, digunakan dalam data mining untuk menemukan hubungan antara item-item dalam dataset [10].

Algoritma ini bekerja dengan mencari frequent itemset, yaitu kumpulan item yang memenuhi nilai minimum support, dan kemudian membangun aturan asosiasi. Itemset dengan frekuensi rendah dihapus. Dua ukuran utama dalam algoritma ini adalah support, yang menunjukkan persentase kombinasi item dalam database, dan confidence, yang menunjukkan kekuatan hubungan antar item. Algoritma berakhir jika

tidak ada itemset umum yang ditemukan dalam satu iterasi.

Langkah dalam analisis pola *frekuensi* tinggi melibatkan pencarian kombinasi item yang memenuhi nilai minimum support yang telah ditetapkan. Dalam konteks ini, probabilitas kemunculan atribut atau sekumpulan atribut A dan B secara bersamaan dalam suatu transaksi disebut support untuk aturan " $A \Rightarrow B$ ". Rumus matematis untuk menghitung nilai support adalah: $\text{Support}(A \Rightarrow B) = P(A \cap B)$, di mana $A \Rightarrow B$ adalah item yang muncul bersamaan, dan $P(A \cap B)$ adalah probabilitas transaksi yang mengandung A dan B, dibagi dengan total jumlah transaksi yang ada. Untuk mendapatkan nilai *support* suatu item, digunakan rumus berikut:

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

Sementara itu, nilai support untuk dua item diperoleh menggunakan rumus berikut:

$$\text{Support}(A, B) = P(A \cap B)$$

$$\text{Support}(A, B) = \frac{\text{Jumlah Transaksi A dan B}}{\text{Total Transaksi}} \times 100\% \quad (2)$$

2.8. Aturan Asosiasi

Menurut Moh. Sholik & Abu Salam, data mining memanfaatkan aturan asosiasi untuk mengidentifikasi hubungan antara item dalam kombinasi tertentu. Dalam teknik ini, terdapat dua parameter utama: *support* dan *confidence*. *Support* menunjukkan *frekuensi* kemunculan itemset dalam seluruh transaksi, sementara *confidence* mengukur kekuatan hubungan antara itemset dalam aturan asosiasi. Untuk menyusun aturan organisasi, penting untuk mencari aturan yang memenuhi *threshold confidence* minimum. *Confidence* dari aturan asosiasi $A \rightarrow B$ dapat dihitung menggunakan rumus berikut:

$$\text{Confidence}(A, B) = \frac{\text{Jumlah Transaksi A dan B}}{\text{Total Transaksi A}} \times 100\% \quad (3)$$

2.9. Lift Ratio

Algoritma Apriori dalam data mining menghasilkan kekuatan hubungan dari kombinasi *itemset* dengan menggunakan alat ukur asosiasi final. Fungsi *lift ratio* adalah untuk mengukur ketepatan dan keakuratan dari alat ukur (*support* dan *confidence*) agar dapat dipercaya sepenuhnya. Cara kerja *lift ratio* yaitu dengan membandingkan antara *confidence* suatu aturan dengan *expected confidence*, yang dapat dihitung dengan rumus berikut:

$$\text{Expected Confidence} = \frac{\sum \text{Transaksi A dan B}}{\sum \text{Transaksi}} \quad (4)$$

Setelah mendapatkan nilai *confidence* dan *expected confidence* dari proses sebelumnya, maka akan diimplementasikan ke dalam rumus *lift ratio* sebagai berikut:

$$\text{Lift Ratio} = \frac{\text{Confidence}}{\text{Expected Confidence}} \quad (5)$$

Jika hasil perhitungan *lift ratio* lebih besar dari 1, maka menunjukkan adanya manfaat dari aturan tersebut. Semakin tinggi nilai *lift ratio*, semakin kuat asosiasinya [11].

2.10. WEKA

WEKA singkatan dari *Waikato Environment for Knowledge Analysis*, ini adalah program berbasis java dan lebih disukai untuk *software machine learning*. WEKA *tools* adalah pemrograman yang dapat diakses secara bebas. Ini mendukung banyak standar penambahan data seperti clustering, preprocessing data, regresi, pemilihan fitur visualisasi, dan klasifikasi. WEKA memungkinkan dalam mencari tahu informasi atau data tersembunyi dari sistem file dan database dengan menggunakan antarmuka visual dan opsi sederhana [12].

3. METODE PENELITIAN

Metodologi yang digunakan dalam penelitian ini adalah Knowledge Discovery in Database (KDD), yang merupakan pendekatan sistematis untuk mengekstraksi pengetahuan dari data besar. Metodologi KDD terdiri dari lima tahapan: Data Selection, Preprocessing (Cleaning), Transformation, Data Mining, dan Evaluation. Tahapan yang dapat diselesaikan dengan metode KDD diantaranya yaitu:

a. Data Selection

Penelitian ini menggunakan data transaksi penjualan produk kemasan dari *platform Shopee*. Data awalnya dalam bentuk Excel dan kemudian masuk ke tahap pemilihan data karena data dari DFS14 masih berupa data mentah. Data yang diambil mencakup transaksi selama 6 bulan pada tahun 2023.

b. Preprocessing atau Cleaning

Setelah tahap data selection selanjutnya membersihkan data awal yang terdiri dari data yang tidak konsisten, kosong, atau identik. Nantinya, sisa data akan diproses untuk menghasilkan data akurat yang diperlukan untuk tahapan selanjutnya.

c. Transformation

Data dimasukkan ke dalam proses pemrosesan data setelah prosedur pembersihan data selesai. Data selanjutnya harus diubah dari format nominal ke format numerik yang dilakukan dengan menggunakan *tools WEKA*.

d. Data Mining

Proses KDD dalam penelitian ini bertujuan untuk menghubungkan keterkaitan antar itemset pada data transaksi produk kemasan menggunakan algoritma Apriori dengan pendekatan aturan asosiasi. Tahapannya meliputi:

- Analisis pola *frekuensi* tinggi
- Pembentukan Aturan Asosiatif
- Aturan Asosiasi Final

e. *Evaluation*

Pada tahap akhir penelitian, dilakukan evaluasi untuk mengukur validitas aturan asosiasi menggunakan metode *lift ratio*. *Lift ratio* dihitung dengan membandingkan nilai *confidence* suatu aturan dengan *expected confidence* sesuai rumus *lift ratio* (5). Hasil perhitungan *lift ratio* yang lebih besar dari 1 menunjukkan manfaat dari aturan tersebut, dengan semakin tinggi nilai *lift ratio* menunjukkan asosiasi yang semakin kuat [11]. Kombinasi itemset dianggap valid jika nilai *lift ratio* >1.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil data transaksi dari penjualan produk kemasan pada toko DFS14 dari bulan Juli-Desember

2023 menghasilkan data sebanyak 552 transaksi, merupakan data mentah yang bersumber dari objek penelitian. Hasil penelitian ini diproses menggunakan metode KDD (*Knowledge Discovery in Database*).

4.2. Data Selection

Pada tahap ini, data awal penjualan sudah berbentuk file .xlsx lalu dilakukan pemilihan data yang dibutuhkan dalam proses selanjutnya, data yang akan diolah hanya beberapa data barang yang mempunyai daya tarik pasar yang lebih banyak. Data yang disalin berupa No. Pesanan, Tanggal, Nama Produk, dan QYT. Dari atribut yang disebutkan, yang akan dipilih untuk proses seleksi yaitu nama barang, tanggal dan jumlah transaksi. Berikut tabel data yang akan diseleksi untuk tahap selanjutnya:

Tabel 1. Contoh Tabel Sebelum Proses Data Selection

No. Pesanan	Tanggal	Box Donat	Box Fc	Box Pizza	Kantong Fc	Kertas Nasi	Lunch Box	Sosis Tray
230701R8MIYQK	01/07/2023	4		4				
230701T2H04VP2	01/07/2023		200		5	10		
230702VWE612DJ	02/07/2023				2	5		
2307031H22NE82	03/07/2023	25			2			1
2307031XR9E97R	03/07/2023		4		1			4
2307032GVP8TA5	03/07/2023		20		1	1		
2307032GVP8TA7	03/07/2023		20		1	1		
23070442JORT1A	04/07/2023				1			2

Berdasarkan tabel 1 diatas atribut yang akan digunakan untuk proses selanjutnya setelah melewati

proses data selection yaitu, nama produk, dan jumlah barang terjual. Data dapat dilihat pada tabel 2.

Tabel 2. Contoh Data Sesudah Proses Data Selection

BOX DONAT	BOX FC	BOX PIZZA	KANTONG FC	KERTAS NASI	LUNCH BOX	SOSIS TRAY
4	200	4	5	10		
			2	5		
25	44		5	2		5
		10	2	2		2
95			41	2		15
30	30	2	4	1	50	8
146			45	1	11	12
...
			5			

Data diatas masih berbentuk data mentah dan belum bisa masuk kedalam proses selanjutnya karena masih terdapat *missing value* yang ditandai dengan adanya kolom yang kosong, maka proses untuk penggantian kolom yang kosong tersebut termasuk kedalam proses selanjutnya yaitu *data cleaning*.

4.3. Preprocessing atau Cleaning

Tahap ini melibatkan pembersihan data (data cleaning) untuk mengatasi nilai kosong (*missing value*) dalam dataset. Dataset harus bebas dari nilai yang hilang, kesalahan, dan outlier sebelum diolah. Jika ada nilai kosong dalam atribut, nilainya diganti dengan angka nol, menunjukkan bahwa tidak ada transaksi yang terjadi untuk item tersebut. Data kolom yang kosong ditampilkan pada Tabel 3.

Tabel 3. Contoh Data Sebelum Proses Preprocessing atau Cleaning

BOX DONAT	BOX FC	BOX PIZZA	KANTONG FC	KERTAS NASI	LUNCH BOX	SOSIS TRAY
4	200	4	5	10		
			2	5		
25	44		5	2		5
		10	2	2		2
95			41	2		15
30	30	2	4	1	50	8
146			45	1	11	12
...
			5			

Selanjutnya yaitu proses data *cleaning* untuk bisa dilakukan proses atau tahap selanjutnya dapat dilihat pada tabel berikut:

Tabel 4. Contoh Data Setelah Proses Preprocessing atau Cleaning

BOX DONAT	BOX FC	BOX PIZZA	KANTONG FC	KERTAS NASI	LUNCH BOX	SOSIS TRAY
4	200	4	5	10	0	0
0	0	0	2	5	0	0
25	44	0	5	2	0	5
0	0	10	2	2	0	2
95	0	0	41	2	0	15
30	30	2	4	1	50	8
146	0	0	45	1	11	12
...
0	0	0	5	0	0	0

Data kosong atau *missing value* yang ditandai dengan kolom yang nilainya hilang kemudian diubah menjadi angka “0”, yang menandakan bahwa barang tersebut tidak ada yang terjual dalam transaksi. Kemudian angka yang berisi angka nominal yaitu barang yang terjual pada data transaksi.

4.4. Transformation

Pada tahap transformation ini akan dilakukan perubahan dataset kedalam data yang dapat dibaca

nantinya oleh aplikasi WEKA. Data akan diubah kedalam bentuk nominal menjadi “Y” dan “N”. Huruf “Y” menandakan bahwa barang tersebut telah terjual, kemudian untuk huruf “N” menandakan bahwa barang tersebut tidak terjual yang sebelumnya ditandai dengan angka “0”. Perubahan data tersebut bertujuan untuk mempermudah melakukan proses data mining. Berikut data yang sudah dirubah menjadi “Y” dan “N” dapat dilihat dalam tabel dibawah ini:

Tabel 5. Contoh Data Sesudah Proses Transformation

BOX DONAT	BOX FC	BOX PIZZA	KANTONG FC	KERTAS NASI	LUNCH BOX	SOSIS TRAY
Y	Y	Y	Y	Y	N	N
N	N	N	Y	Y	N	N
Y	Y	N	Y	Y	N	Y
N	N	Y	Y	Y	N	Y
Y	N	N	Y	Y	N	Y
Y	Y	Y	Y	Y	Y	Y
Y	N	N	Y	Y	Y	Y
...
N	N	N	Y	N	N	N

Seperti yang sudah dijelaskan sebelumnya dan ditunjukkan oleh tabel yang sudah diubah pada dataset, bahwa huruf “Y” bisa dikatakan “1” dan huruf “N” bisa disebut angka “0”. Selanjutnya, pada tahap transformasi data, dilakukan juga perubahan format file dataset dari Excel, dari format .xlsx ke .csv, dengan tujuan mempermudah proses data mining yang akan dilakukan menggunakan *tools* WEKA 3.8.

4.5. Data Mining

Pada tahap proses data mining dilakukan 2 teknis pengujian yaitu perhitungan menggunakan *tools* dan perhitungan manual. Dilakukan beberapa percobaan terlebih dahulu menggunakan *tools* untuk mendapatkan pola yang tidak terlalu banyak namun optimal, maka ditentukan bahwa minimum nilai support sebesar 40% dan minimum nilai confidence sebesar 80%. Menurut Kusumo, semakin tinggi

minimal support dan minimal confidence yang digunakan maka aturan yang terbentuk akan semakin sedikit dan menurun akurasi [7]. Tahapan data mining

4.6. Analisis Pola Frekuensi Tinggi

Pada tahap ini akan dilakukan pencarian kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam database. Nilai *support* sebuah item diperoleh dengan rumus persamaan (1). Dengan contoh perhitungan:

$$\text{Support (A)} = \frac{160}{184} = 0,86957$$

Persamaan diatas dilakukan terus untuk mencari nilai *support* dari setiap item yang ada di dalam dataset, berikut hasil semua perhitungan yang sesuai dengan persamaan mencari nilai *support*:

Tabel 6. Hasil Perhitungan *support* 1 Itemset

No	ITEMSET	SUPPORT COUNT	SUPPORT
1	Kantong Fc	160	0,86957
2	Sosistray	125	0,67935
3	Box Donat	118	0,64130
4	Kertas Nasi	114	0,61957
5	Box Fc	89	0,48370
6	Lunch Box	69	0,37500
7	Box Pizza	56	0,30435

Tabel tersebut menunjukkan perhitungan nilai *support* dari dataset. Langkah selanjutnya adalah mencari nilai *support* untuk 2 itemset dengan menggabungkan itemset dari tahap sebelumnya dan memindai ulang database. Itemset yang memenuhi minimum *support* 40% dipilih sebagai pola frequent tinggi. Tabel tersebut menampilkan itemset yang memenuhi kriteria ini.

Tabel 7. Nilai *Support* diatas 40%

No	ITEMSET	SUPPORT COUNT	SUPPORT
1	Kantong Fc	160	0,86957
2	Sosistray	125	0,67935
3	Box Donat	118	0,64130
4	Kertas Nasi	114	0,61957
5	Box Fc	89	0,48370

Setelah mengetahui itemset yang memenuhi minimum nilai *support* akan dipilih sebagai pola *frequent* tinggi dari kandidat, selanjutnya adalah mencari kombinasi 2 itemset, dengan rumus persamaan (2). Dengan contoh perhitungan:

$$\text{Support (A} \cap \text{B)} = \frac{98}{184} = 0,533$$

Persamaan untuk menghitung 2 kombinasi itemset juga perlu dilakukan dengan semua item yang ada di dalam database, berikut adalah hasil perhitungan dengan menggunakan 2 itemset:

Tabel 8. Contoh Hasil Perhitungan *support* 2 Itemset

ITEMSET	JUMLAH		SUPPORT
Box Donat , Kantong Fc	98	184	0,533
Box Donat , Sosistray	87	184	0,473
Box Fc , Kantong Fc	84	184	0,457
Kantong Fc , Box Donat	98	184	0,533
Kantong Fc , Box Fc	84	184	0,457
Kantong Fc , Kertas Nasi	102	184	0,554

Dari perhitungan diatas, masih ada kemungkinan untuk melakukan perhitungan sampai dengan 3 itemset, maka dilakukan perhitungan dengan persamaan atau rumus yang sama untuk mencari 2 itemset sebelumnya. selanjutnya adalah mencari kombinasi 3 itemset, dengan rumus persamaan (2.1). Berikut adalah contoh perhitungan *support* 3 itemset.

$$\text{Support (A} \cap \text{B)} = \frac{27}{184} = 0,147$$

Perhitungan diatas juga akan dilakukan kembali dengan menghitung semua kombinasi dari item-item yang telah didapatkan dari perhitungan itemset. Berikut adalah tabel hasil yang diperoleh dari perhitungan untuk mencari 3 itemset:

Tabel 9. Contoh Hasil Perhitungan *Support* 3 Itemset

ITEMSET	JUMLAH		SUPPORT
Box Donat, Box Fc, Box Pizza	27	184	0,147
Box Donat, Box Fc, Kantong Fc	53	184	0,288
Box Donat, Box Fc, Kertas Nasi	40	184	0,217
Box Donat, Box Fc, Lunch Box	28	184	0,152
Box Donat, Box Fc, Sosistray	42	184	0,228
Box Donat, Box Pizza, Kantong Fc	34	184	0,185

Perhitungan dilakukan secara bertahap untuk menentukan nilai *support* dengan menggabungkan itemset, dimulai dari 2 itemset hingga 3 itemset. Jika *support* dari setiap itemset masih mencukupi, pencarian dilanjutkan dengan menggabungkan lebih banyak itemset. Namun, karena tidak memungkinkan untuk menggabungkan 4 itemset, pencarian dihentikan pada tahap 3 itemset.

4.7. Pembentukan Aturan Asosiatif

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence*, dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus (3).

$$\text{Confidence} = P(B|A) = \frac{\text{Jumlah Transaksi A dan B}}{\text{Jumlah Transaksi A}}$$

Sebelum memulai untuk mencari aturan asosiatif tentukan terlebih dahulu nilai minimum *confidence* yang akan ditentukan yaitu sebesar 80% atau 0.8. Jadi pada saat perhitungan berjalan apabila hasil dari nilai *confidence* tiap itemset kurang dari 80% maka itemset tersebut tidak akan digunakan untuk ke tahap

selanjutnya. Pada tahap pencarian *confidence* dengan rumus dan persamaan yang sudah dilakukan sebelumnya, maka perhitungan dimulai dari kombinasi 2 itemset. Selanjutnya mencari nilai *confidence* dari hasil pencarian minimum *support*, dengan contoh perhitungan:

$$Confidence = P(B|A) = \frac{57}{118} = 48,305\%$$

Untuk mencari keseluruhan nilai *confidence* dari masing-masing itemset, dilakukan dengan cara yang sama seperti cara diatas dengan rumus dan persamaan yang sama juga. Tabel berikut merupakan hasil dari perhitungan nilai *confidence*:

Tabel 10. Contoh Hasil Perhitungan *Confidence* 2 Itemset

ITEMSET	SUPPORT COUNT (A∩B)	Transaksi A	CONFIDENCE
Box Donat => Box Fc	57	118	48,305%
Box Donat => Box Pizza	39	118	33,051%
Box Donat => Kantong Fc	98	118	83,051%
Box Donat => Kertas Nasi	71	118	60,169%
Box Donat => Lunch Box	55	118	46,610%
Box Donat => Sosistray	87	118	73,729%

Tahapan selanjutnya setelah mendapatkan hasil *confidence* dengan menggunakan kombisanasi 2 itemset, ialah melakukan perhitungan untuk mencari *confidence* dari kombinasi 3 itemset. Untuk perhitungan *confidence* kombinasi 3 itemset masih dilakukan dengan persamaan yang sama untuk mencari nilai *confidence* 2 itemset. Berikut ini hasil dari perhitungan *confidence* dari kombinasi 3 itemset.

Tabel 11. Contoh Hasil Perhitungan *Confidence* 3 Itemset

ITEMSET	SUPPORT COUNT (A∩B)	Transaksi A	CONFIDENCE
Box Donat, Box Pizza => Kertas Nasi	56	118	47,458%
Box Donat, Kantong Fc => Kertas Nasi	62	118	52,542%
Box Donat, Kantong Fc => Sosistray	74	118	62,712%
Box Donat, Kertas Nasi => Sosistray	56	118	47,458%
Box Fc, Kantong Fc => Kertas Nasi	56	89	62,921%
Box Fc, Kantong Fc => Sosistray	51	89	57,303%
Box Pizza, Kantong Fc => Kertas Nasi	35	56	62,500%

ITEMSET	SUPPORT COUNT (A∩B)	Transaksi A	CONFIDENCE
Box Pizza, Kantong Fc => Sosistray	32	56	57,143%
Box Pizza, Kertas Nasi => Sosistray	27	56	48,214%
Kantong Fc, Kertas Nasi => Sosistray	76	160	47,500%

Setelah menghitung dan mencari nilai minimum *support* dan *confidence*, selnjutnya membuat beberapa rule atau aturan dari itemset yang sudah memenuhi ketentuan nilai minimum *confidence* yang sudah ditentukan yaitu 80%.

4.8. Aturan Asosiasi Final

Rule atau aturan asosiatif yang sudah didapat dari hasil penghitungan nilai minimum *support* dan minimum *confidence* ada pada tabel dibawah ini:

Tabel 12. Aturan Asosiatif

No	ITEMSET	SUPPORT	CONFIDENCE
1	Box Fc => Kantong Fc	45,652%	94,382%
2	Lunch Box => Kantong Fc	34,008%	90,688%
3	Kertas Nasi => Kantong Fc	55,435%	89,474%
4	Kertas Nasi, Sosistray => Kantong Fc	55,435%	89,474%
5	Box Pizza => Kantong Fc	27,174%	89,286%
6	Box Pizza, Lunch Box => Kantong Fc	27,174%	89,286%
7	Sosistray => Kantong Fc	59,239%	87,200%
8	Box Pizza, Sosistray => Kantong Fc	26,174%	86,000%
9	Box Donat, Sosistray => Kantong Fc	53,300%	85,113%
10	Box Donat => Kantong Fc	53,261%	83,051%
11	Lunch Box => Kertas Nasi	30,978%	82,609%

Hasil dari perhitungan diatas terbentuk 11 aturan asosiatif berdasarkan parameter yang telah ditentukan sebelumnya yaitu, dengan minimum nilai *support* 40% dan minimum nilai *confidence* 80%.

4.9. Evaluation

Tahapan terakhir dari proses KDD adalah evaluasi, dimana pada tahap ini menentukan pada aturan asosiasi di atas dapat dinyatakan valid. Pada penelitian ini akan menggunakan sebuah metode *lift ratio* dengan perhitungan sebagai berikut:

- Mencari nilai *Expected Confidence* menggunakan rumus persamaan (3). Hasil perhitungannya dapat dilihat pada tabel 13 dibawah ini:

Tabel 13. Hasil Perhitungan *Expected Confidence*

No	Itemset	Transaksi A	Total Transaksi	EC
1	Box Fc , Kantong Fc	89	184	48,37%
2	Lunch Box , Kantong Fc	115	184	62,50%
3	Kertas Nasi , Kantong Fc	114	184	61,96%
4	Kertas Nasi , Sosistray , Kantong Fc	85	184	46,20%
5	Box Pizza , Kantong Fc	128	184	69,57%
6	Box Pizza, Lunch Box , Kantong Fc	84	184	45,65%
7	Sosistray , Kantong Fc	125	184	67,93%
8	Box Pizza, Sosistray , Kantong Fc	90	184	48,91%
9	Box Donat, Sosistray , Kantong Fc	87	184	47,28%
10	Box Donat , Kantong Fc	118	184	64,13%
11	Lunch Box , Kertas Nasi	57	184	30,98%

- b. Mencari nilai Lift Ratio dengan rumus yang ada pada (2.5)

Hasil perhitungannya dapat dilihat pada tabel 4.14 dibawah ini:

Dengan rumus yang ada pada (5) Hasil perhitungannya dapat dilihat pada tabel 14 dibawah ini:

Tabel 14. Hasil Perhitungan *Lift Ratio*

No	Itemset	Confidence	EC	Lift ratio
1	Box Fc, Kantong Fc	94,38%	48,37%	1,951
2	Lunch Box, Kantong Fc	90,69%	62,50%	1,451
3	Kertas Nasi, Kantong Fc	89,47%	61,96%	1,444
4	Kertas Nasi, Sosistray, Kantong Fc	89,47%	46,20%	1,937
5	Box Pizza, Kantong Fc	89,29%	69,57%	1,283
6	Box Pizza, Lunch Box, Kantong Fc	89,29%	45,65%	1,956
7	Sosistray, Kantong Fc	87,20%	67,93%	1,284
8	Box Pizza, Sosistray, Kantong Fc	86,00%	48,91%	1,758
9	Box Donat, Sosistray, Kantong Fc	85,11%	47,28%	1,800
10	Box Donat, Kantong Fc	83,05%	64,13%	1,295
11	Lunch Box, Kertas Nasi	82,61%	30,98%	1,667

Pada tahap evaluasi ini menguji hasil pembentukan aturan asosiasi dengan pola aturan yang terbentuk sebanyak 11 *rule* dengan hasil nilai *lift ratio* >1.0 maka pola yang telah terbentuk dikatakan valid dan sudah dapat digunakan untuk meningkatkan strategi penjualan di DFS14.

4.10. Pembahasan

Penelitian dengan algoritma apriori dan metodologi KDD menggunakan tools WEKA pada 552 transaksi menemukan 11 aturan terbaik dalam penjualan produk kemasan, dengan nilai minimum

support 40% dan confidence 80%. Nilai minimum support 40% memastikan aturan hanya mencakup item yang muncul di setidaknya 40% transaksi, sementara confidence 80% menunjukkan tingkat kepastian tinggi bahwa item terkait juga akan dibeli. Semua aturan diuji dengan evaluasi lift ratio dan valid jika nilai lift ratio >1, yang menunjukkan manfaat dari aturan tersebut. Lima aturan teratas dengan nilai *confidence* dan *lift ratio* terbesar ditampilkan pada tabel 15.

Tabel 15. Hasil Evaluasi *Lift Ratio* Tertinggi

No	Itemset	Confidence	EC	Lift ratio
1	Box Pizza, Lunch Box, Kantong Fc	0,8929	0,4565	1,956
2	Box Fc , Kantong Fc	0,9438	0,4837	1,951
3	Kertas Nasi, Sosistray, Kantong Fc	0,8947	0,4620	1,937
4	Box Donat, Sosistray, Kantong Fc	0,8511	0,4728	1,800
5	Box Pizza, Sosistray, Kantong Fc	0,8600	0,4891	1,758

Hasil analisis menunjukkan bahwa produk KANTONG FC sering dibeli bersama dengan BOX FC, KERTAS NASI, dan SOSISTRAY. Ini menandakan bahwa KANTONG FC adalah produk penting yang biasanya dibeli bersama produk kemasan makanan lainnya. Oleh karena itu, jika pelanggan membeli salah satu dari produk ini, mereka kemungkinan besar juga akan membeli KANTONG FC. Berdasarkan penelitian ini, toko dapat membuat penawaran *bundling* untuk produk-produk tersebut, seperti:

- 4.1. KANTONG FC dengan BOX FC
- 4.2. KANTONG FC dengan KERTAS NASI
- 4.3. KANTONG FC dengan SOSISTRAY

Toko dapat menawarkan diskon atau harga paket khusus untuk pembelian bundling ini. Strategi ini dapat menarik konsumen karena memudahkan mereka mendapatkan produk yang sering dibeli bersama dalam satu transaksi.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang menggunakan algoritma Apriori dengan tools WEKA dan metode KDD pada 552 transaksi selama 6 bulan, ditemukan 11 aturan terbaik dengan nilai minimum *support* 40% dan *confidence* 80%. Hasil analisis menunjukkan bahwa nilai *lift ratio* yang lebih dari 1,0 mengindikasikan bahwa pola-pola ini valid dan dapat digunakan untuk meningkatkan strategi penjualan di DFS14, dengan membantu menilai kekuatan hubungan antara produk dalam pola pembelian dan memastikan bahwa pembelian bersama produk tidak terjadi secara kebetulan.

Penelitian juga mengungkap bahwa produk KANTONG FC sering dibeli bersamaan dengan BOX FC, KERTAS NASI, dan SOSISTRAY. Toko dapat memanfaatkan informasi ini untuk meningkatkan penjualan dengan menawarkan diskon atau harga

paket khusus untuk kombinasi produk tersebut, yang juga mempermudah konsumen dalam melakukan pembelian sekaligus. Untuk penelitian selanjutnya, disarankan agar peneliti menggunakan tools yang berbeda, objek yang berbeda, dan dataset yang lebih besar untuk mendapatkan hasil yang lebih representatif. Selain itu, variasi nilai support dan confidence yang lebih luas dapat digunakan untuk memperoleh asosiasi rule yang lebih banyak. Penelitian selanjutnya juga dapat dikembangkan dengan menerapkan metode algoritma lain, seperti FP-Growth atau Hash Based, untuk memperoleh hasil yang lebih komprehensif dan mendalam.

DAFTAR PUSTAKA

- [1] & C. M. A. Monavia Ayu Rizaty, "Data GMV E-Commerce di Asia Tenggara pada 2020 hingga 2023," *Dataindonesia.id*, 2023. website: <https://dataindonesia.id/ekonomi-digital/detail/data-gmv-ecommerce-di-asia-tenggara-pada-2020-hingga-2023%0A%0A>
- [2] C. Kleisiari, M. N. Duquenne, and G. Vlontzos, "E-commerce in the retail chain store market: An alternative or a main trend?," *Sustain.*, vol. 13, no. 8, 2021, doi: 10.3390/su13084392.
- [3] M. Rizki, D. Devrika, I. H. Umam, and F. S. Lubis, "Aplikasi Data Mining dalam Penentuan Layout Swalayan dengan Menggunakan Metode MBA," *J. Tek. Ind. J. Has. Penelit. dan Karya Ilm. dalam Bid. Tek. Ind.*, vol. 5, no. 2, p. 130, 2020, doi: 10.24014/jti.v5i2.8958.
- [4] M. Hafizh, T. Novita, D. Guswandi, H. Syahputra, and L. Mayola, "Implementasi Data Mining Menggunakan Algoritma Fp-Growth Untuk Menganalisa Transaksi Penjualan Ekspor Online," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 3, pp. 242–249, 2023, doi: 10.47233/jteksis.v5i3.847.
- [5] F. S. Amalia and D. D. Setiawansyah, "Analisis Data Penjualan Handphone Dan Elektronik Menggunakan Algoritma Apriori (Studi Kasus: Cv Rey Gasendra)," ... *J. Telemat.* ..., vol. 2, no. 1, pp. 1–6, 2021, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/telefortech/article/view/1810>
- [6] A. B. Setiawan, D. P. Pamungkas, T. Krisna, and W. Aprilia, "PENGELOMPOKAN MERK OLI MENGGUNAKAN K-MEANS PADA TOKO," vol. 4, no. 1, 2023.
- [7] H. Kusumo, E. Sedyono, and M. Marwata, "Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 49, 2019, doi: 10.21580/wjit.2019.1.1.4000.
- [8] G. Thabroni, "Kemasan Produk: Pengertian, Fungsi, Tujuan, Dimensi, Merancang, dsb," *serupa.id*, 2022. <https://serupa.id/kemasan-produk-pengertian-fungsi-tujuan-dimensi-merancang-dsb/>
- [9] S. Manaf, P. P. Makassar, L. D. Hasan, M. Musawantoro, F. A. Zaenal, and P. P. Makassar, "Pengaruh Kegiatan Pemasaran Online (E-Commerce) terhadap Keunggulan Bersaing Produk Kuliner Tradisional Di Sulawesi Selatan Effect of Online Marketing (E-Commerce) Activities," no. October, 2020, doi: 10.33649/pusaka.v2i1.47.
- [10] E. Alma'arif, E. Utami, and F. W. Wibowo, "Implementasi Algoritma Apriori Untuk Rekomendasi Produk Pada Toko Online," *Creat. Inf. Technol. J.*, vol. 7, no. 1, p. 63, 2021, doi: 10.24076/citec.2020v7i1.241.
- [11] I. M. D. P. Asana, I. G. I. Sudipa, A. A. T. W. Mayun, N. P. S. Meinarni, and D. V. Waas, "Aplikasi Data Mining Asosiasi Barang Menggunakan Algoritma Apriori-TID," *INFORMAL Informatics J.*, vol. 7, no. 1, p. 38, 2022, doi: 10.19184/isj.v7i1.30901.
- [12] A. Pangestu and T. Ridwan, "Penerapan Data Mining Menggunakan Algoritma K-Means Pengelompokan Pelanggan Berdasarkan Kubikasi Air Terjual Menggunakan Weka," *JUST IT J. Sist. Informasi, Teknol. Inf. dan Komput.*, vol. 11, no. 3, pp. 67–71, 2022, [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/article/view/11591>