



Implementasi Algoritma Apriori dalam Menemukan Pola Asosiasi pada Data Penjualan Produk Retail

Sufajar Butsianto*, Candra Naya, Anggi Muhammad Rifa'i

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

Email: ¹*sufajar@pelitabangsa.ac.id, ²candranaya@pelitabangsa.ac.id, ³anggi@pelitabangsa.ac.id

Email Penulis Korespondensi: sufajar@pelitabangsa.ac.id

Abstrak—Penelitian ini bertujuan untuk mengimplementasikan algoritma Apriori dalam menemukan pola asosiasi pada data penjualan produk retail, dengan menggunakan pendekatan Association Rule Mining. Mengevaluasi ruler atau aturan asosiasi yang terbentuk berdasarkan parameter support, confidence, dan lift, dalam menemukan pola asosiasi pada data penjualan produk retail dengan fokus pada keterkaitan antar kategori produk. Data yang digunakan terdiri dari 500 data penjualan sebagai data sampel dan 5.972 transaksi sebagai data uji. Proses penambangan data dilakukan terhadap kategori produk utama seperti Milk, Coffee, Tea, Drinks, Detergent, dan Biscuit/Snacks, guna menemukan aturan asosiasi yang muncul secara bersamaan dengan kategori Bulk Products dalam satu waktu transaksi. Parameter minimum support ditetapkan sebesar 0,02 dan minimum confidence sebesar 0,5. Dengan menggunakan parameter tersebut, diperoleh beberapa aturan asosiasi yang signifikan. Salah satu aturan terkuat menunjukkan bahwa apabila produk dalam kategori Milk, Coffee, Tea, Drinks, Detergent, dan Biscuit/Snacks dibeli secara bersamaan, maka terdapat kemungkinan sebesar 64,3% (confidence) bahwa produk dalam kategori Bulk Products juga dibeli pada waktu yang sama. Nilai support dari aturan ini mencapai 3,8%, dan lift sebesar 1,49, yang menunjukkan adanya hubungan asosiasi positif dan tidak terjadi secara kebetulan. Evaluasi terhadap data uji menunjukkan bahwa pola tersebut konsisten ditemukan dalam 5.972 transaksi, dengan tingkat keberulangan sebesar 61,7%. Hasil penelitian ini menunjukkan bahwa algoritma Apriori efektif dalam mengidentifikasi pola pembelian konsumen yang dapat dimanfaatkan untuk strategi penempatan produk, penawaran bundling, dan perencanaan stok dalam manajemen ritel.

Kata kunci: Algoritma Apriori; Association Rule Mining; Confidence; Lift; Kategori Produk; Bulk Products; Penjualan Retail

Abstract—This study aims to implement the Apriori algorithm in finding association patterns in retail product sales data, using the Association Rule Mining approach. Evaluating the ruler or association rules formed based on the support, confidence, and lift parameters, in finding association patterns in retail product sales data with a focus on the relationship between product categories. The data used consists of 500 sales data as sample data and 5,972 transactions as test data. The data mining process was carried out on the main product categories such as Milk, Coffee, Tea, Drinks, Detergent, and Biscuit/Snacks, to find association rules that appear simultaneously with the Bulk Products category in one transaction time. The minimum support parameter was set at 0.02 and the minimum confidence was set at 0.5. By using these parameters, several significant association rules were obtained. One of the strongest rules shows that if products in the Milk, Coffee, Tea, Drinks, Detergent, and Biscuit/Snacks categories are purchased together, then there is a 64.3% probability (confidence) that products in the Bulk Products category are also purchased at the same time. The support value of this rule reached 3.8%, and the lift value was 1.49, indicating a positive association and not a coincidence. Evaluation of the test data showed that this pattern was consistently found across 5,972 transactions, with a repeatability rate of 61.7%. The results of this study demonstrate that the Apriori algorithm is effective in identifying consumer purchasing patterns that can be utilized for product placement strategies, bundling offers, and inventory planning in retail management.

Keywords: Apriori Algorithm; Association Rule Mining; Confidence; Lift; Bulk Products; Retail Sales Data; Retail Strategy

1. PENDAHULUAN

Dalam era digital dan persaingan bisnis yang semakin ketat[1], perusahaan retail dituntut untuk mampu memahami perilaku konsumen secara lebih mendalam guna meningkatkan efisiensi strategi pemasaran dan penjualan[2]. Salah satu aset berharga yang dimiliki oleh perusahaan retail adalah data transaksi penjualan yang tersimpan dalam jumlah besar[3]. Meskipun data penjualan produk retail yang dikumpulkan dalam jumlah besar dapat mencerminkan aktivitas bisnis yang kompleks dan dinamis, data mentah pada dasarnya tidak memiliki makna yang signifikan apabila tidak diolah dan dianalisis dengan metode yang tepat. Dalam konteks penelitian ini, data penjualan yang terdiri atas ribuan transaksi hanya akan menjadi sekumpulan angka dan kategori produk tanpa nilai informasi apabila tidak dilakukan analisis lebih lanjut. Oleh karena itu, diperlukan pendekatan yang sistematis seperti algoritma Apriori untuk menggali pola-pola tersembunyi di balik data tersebut. Dengan menerapkan algoritma Apriori, data mentah tersebut dapat dianalisis untuk menemukan asosiasi atau keterkaitan antarproduk dalam satu transaksi. Pola-pola ini, yang diukur melalui parameter seperti support, confidence, dan lift, memberikan wawasan yang bernilai bagi pengambilan keputusan bisnis, misalnya dalam perencanaan promosi produk, penataan layout toko, dan strategi penjualan silang (cross-selling). Dengan demikian, penelitian ini menegaskan bahwa nilai sesungguhnya dari data terletak pada kemampuannya untuk diubah menjadi informasi yang dapat ditindaklanjuti melalui analisis yang tepat.

Di sinilah peran data mining menjadi sangat krusial, khususnya dalam mengidentifikasi pola atau hubungan tersembunyi antar produk yang sering dibeli secara bersamaan[4]. Ritel merupakan salah satu cara pemasaran produk meliputi semua aktivitas yang melibatkan penjualan barang secara langsung ke konsumen akhir untuk penggunaan pribadi dan bukan bisnis[5]. Organisasi ataupun seseorang yang menjalankan bisnis ini disebut pula sebagai pengecer[6]. Pada praktiknya pengecer melakukan pembelian barang ataupun produk dalam jumlah besar dari produsen, ataupun pengimport baik secara langsung ataupun melalui grosir, untuk kemudian dijual kembali dalam jumlah kecil[7]. Saat ini memiliki masalah shrinkage[8] (barang tidak terjual, expired dan barang terbuang) yang cenderung naik setiap bulannya dan



melewati budget konsekuensi yang telah ditentukan oleh pusat.[9], [10]. pemanfaatan data operasional yang ada untuk menunjang kegiatan pengambilan keputusan dan diperlukan suatu analisis data untuk menggali potensi-potensi informasi yang ada, berusaha untuk memanfaatkan data yang sudah dimiliki untuk menggali informasi yang berguna membantu mengambil keputusan.[11] Penggunaan teknik data mining diharapkan dapat memberikan pengetahuan-pengetahuan yang sebelumnya tersembunyi di dalam data sehingga menjadi informasi yang berharga[12]. Data mining[13] memainkan peran penting sebagai suatu pendekatan analitis untuk menggali informasi tersembunyi dari kumpulan data yang besar. Salah satu teknik dalam data mining yang banyak digunakan dalam bidang penjualan adalah association rule mining[14], yaitu metode yang bertujuan untuk menemukan aturan asosiasi atau hubungan antar item dalam sebuah transaksi. Dalam konteks ini, penting untuk mengidentifikasi keterkaitan antar kategori produk guna mengetahui kemungkinan terjadinya pembelian bersamaan.

Pencarian pola penjualan berawal dari pengolahan data barang dari setiap pembelian[15]. Algoritma apriori berfungsi untuk membentuk kandidat kombinasi item yang mungkin, lalu diuji apakah kombinasi tersebut memenuhi parameter support dan confidence minimum[16]. Salah satu teknik data mining yang efektif untuk menggali pola pembelian konsumen adalah association rule mining[17]. Algoritma Apriori merupakan salah satu algoritma paling populer dalam association rule mining yang bekerja dengan mencari frequent itemset dan menghasilkan aturan asosiasi berdasarkan nilai support dan confidence[18].

Dengan menerapkan algoritma Apriori[19], penelitian ini bertujuan untuk memberikan pemahaman yang lebih mendalam tentang pola penjualan produk di retail, serta memberikan rekomendasi yang dapat meningkatkan penjualan dan kepuasan pelanggan. Selain itu, hasil dari penelitian ini diharapkan dapat menjadi acuan bagi para pelaku industri retail dalam mengambil keputusan yang berbasis data, sehingga dapat bersaing lebih efektif di pasar yang semakin kompetitif. Algoritma Apriori memiliki keunggulan dalam menyaring kombinasi item yang sering muncul dalam transaksi[20], namun juga menghadapi tantangan dalam hal efisiensi ketika diterapkan pada dataset yang sangat besar[20]. Oleh karena itu, implementasi algoritma ini perlu disesuaikan dengan kebutuhan serta karakteristik data yang digunakan. Pada penelitian ini, algoritma Apriori diimplementasikan untuk menganalisis pola penjualan produk pada sebuah toko retail, dengan tujuan untuk mengidentifikasi kombinasi produk yang sering dibeli bersamaan oleh konsumen[21]. Penelitian ini tidak hanya bertujuan untuk mengekstrak informasi tersembunyi dari data penjualan, tetapi juga sebagai upaya memberikan rekomendasi strategis berbasis data guna mendukung pengambilan keputusan bisnis. Dengan demikian, hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan efektivitas pengelolaan produk, penataan toko, serta strategi pemasaran pada sektor retail[22].

Meskipun algoritma Apriori telah banyak diterapkan dalam analisis data penjualan untuk menemukan pola asosiasi antar produk[23], sebagian besar penelitian sebelumnya cenderung berfokus pada identifikasi aturan asosiasi secara umum tanpa mempertimbangkan kedalaman konteks kategori produk yang spesifik dan relevansi temporal[24] (waktu transaksi bersamaan). Selain itu, sebagian besar studi menggunakan data simulasi atau dataset yang bersifat publik dan terbatas dalam merepresentasikan kompleksitas transaksi retail yang sebenarnya. Hal ini menimbulkan kesenjangan penelitian (research gap) dalam penerapan algoritma Apriori yang mampu mengungkap korelasi kuat antara kategori produk tertentu dengan akurasi dan evaluasi berbasis metrik seperti support, confidence, dan lift yang terukur secara persentase[25].

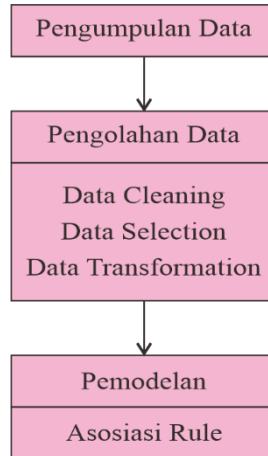
Penelitian ini hadir untuk mengisi celah tersebut dengan pendekatan yang lebih terarah dan kontekstual, yakni dengan menggunakan data nyata berupa 500 data penjualan sebagai sampel dan 5.972 transaksi sebagai data uji, serta memfokuskan analisis pada keterkaitan antara kategori produk Milk, Coffee, Tea, Drinks, Detergent, dan Biscuit/Snacks terhadap kemunculan produk dalam kategori Bulk Products dalam satu waktu transaksi. Temuan penelitian menunjukkan bahwa apabila produk-produk dalam kategori tersebut dibeli bersamaan, maka terdapat kemungkinan sebesar 64,3% (confidence) bahwa produk dari kategori Bulk Products juga dibeli, dengan nilai support sebesar 3,8% dan lift sebesar 1,49. Evaluasi pada data uji memperkuat temuan ini dengan tingkat keberulangan sebesar 61,7%, yang menunjukkan konsistensi pola asosiasi pada skala data yang lebih besar. Kebaruan dari penelitian ini terletak pada pendekatan kombinatorik antara segmentasi kategori produk secara tematik dan analisis waktu transaksi bersamaan, yang belum banyak dibahas dalam penelitian terdahulu. Selain itu, evaluasi terhadap rule yang terbentuk tidak hanya dilakukan pada satu dataset, tetapi juga diuji validitasnya pada data yang lebih besar, sehingga hasilnya lebih dapat diandalkan dan aplikatif dalam konteks ritel modern.

Penelitian ini bertujuan untuk mengimplementasikan algoritma Apriori dalam proses analisis data penjualan retail[26] guna menemukan pola asosiasi antar kategori produk yang sering muncul secara bersamaan, mengidentifikasi keterkaitan spesifik antara produk dalam kategori Milk, Coffee, Tea, Drinks, Detergent, dan Biscuit/Snacks dengan kategori Bulk Products dalam satu waktu transaksi, mengevaluasi kekuatan dan validitas aturan asosiasi (rule) berdasarkan parameter support, confidence, dan lift, serta menguji konsistensinya menggunakan data uji sebanyak 5.972 transaksi, memberikan dasar pengambilan keputusan yang berbasis data untuk strategi penempatan produk, perencanaan promosi bundling, dan pengelolaan stok di sektor ritel. Penelitian ini memberikan kontribusi dan manfaat seperti memberikan pendekatan evaluatif yang lebih spesifik terhadap association rule mining dengan mempertimbangkan konteks kategori produk dan transaksi simultan, serta pengujian terhadap dataset yang lebih besar untuk mengukur generalisasi pola. Serta memberikan insight berbasis data kepada pelaku usaha retail mengenai kecenderungan konsumen dalam membeli produk secara bersamaan, khususnya keterkaitan kategori utama dengan produk Bulk, sehingga dapat dioptimalkan dalam strategi pemasaran, penyusunan display, dan pengendalian logistik.



2. METODOLOGI PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan yang sistematis agar tujuan penelitian dapat tercapai dengan optimal. Tahapan-tahapan tersebut dijelaskan Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Dalam penelitian ini data yang digunakan menggunakan data pembelian dari PT. XYZ. Penggunaan data pembelian dari PT. XYZ dianggap relevan karena mampu menggambarkan perilaku konsumen dalam konteks nyata, sehingga hasil analisis diharapkan dapat memberikan rekomendasi strategis bagi perusahaan, terutama dalam hal peningkatan penjualan dan pengelolaan persediaan barang. Data yang digunakan adalah data dari bulan September 2023 sampai bulan Februari 2024 (6 bulan terakhir) sebanyak 5972 data yang memiliki nilai total penjualan diatas rata-rata yaitu lebih besar dari Rp. 1.435.010. Dalam Tabel 1 adalah data sampel penjualan.

Tabel 1. Data Sampel Penjualan

Keycode Trans	Cat	Sum of Sale amt
2023/09/02-06090003170200	Bulk Product	Rp 299.400.000
2023/09/03-06090003170200	Bulk Product	Rp 252.000.000
2023/09/06-06090003168550	Bulk Product	Rp 567.000.000
2023/09/09-06090000685455	Bulk Product	Rp 310.320.000
2023/09/09-06090003168550	Bulk Product	Rp 554.400.000
2023/09/10-06090000685476	Bulk Product	Rp 189.000.000
2023/09/18-06090003170200	Bulk Product	Rp 182.275.000
2023/09/18-06090003987855	Bulk Product	Rp 236.000.000
2023/09/19-06090003168550	Bulk Product	Rp 227.960.000
2023/09/24-06005000003439	Bulk Product	Rp 638.000.000
2023/09/25-06090001515418	Bulk Product	Rp 319.000.000
2023/09/25-06090004033432	Bulk Product	Rp 176.400.000
2023/09/26-06090004033432	Bulk Product	Rp 295.500.000
2023/09/27-06090003168550	Bulk Product	Rp 1.382.962.280
.....		

Agar penyajian data lebih mudah maka dibuat format biner 1/0 dalam tabular form, representasi biner 0/1 menggambarkan bahwa kategori produk yang bernilai 1 berarti item dalam kategori tersebut dibeli oleh konsumen, sedangkan kategori produk yang bernilai 0 atau kosong berarti item dalam kategori tersebut tidak dibeli konsumen. Setelah format data transaksi diubah menjadi data transaksi dalam representasi biner 0/1 kemudian data set transaksi yang disimpan dalam .xls. Data transaksi dalam representasi biner 0/1 terdiri dari baris yang menunjukkan transaksi dan kolom yang menunjukkan kategori produk.

Tabel 2 menjadi representasi biner 0/1 pada setiap item yang dibeli pada setiap transaksi ditandai dengan angka 1 pada kategori produk yang memuat item tersebut.

Tabel 2. Data Transaksi dalam Representasi Biner 0/1

Keycode Trans	APPLIANCE	Bakery	Biscuit/Snacks
2023/09/01-06008000315258	0	0	1
2023/09/01-06090000566687	0	0	1
2023/09/01-06090000567854	1	0	1



Keycode Trans	APPLIANCE	Bakery	Biscuit/Snacks
2023/09/01-06090000567912	0	0	1
2023/09/01-06090000567971	0	0	1
2023/09/01-06090000570528	0	0	1
2023/09/01-06090000573793	0	0	1
2023/09/01-06090000598282	0	0	0
2023/09/01-06090000605104	0	0	0
2023/09/01-06090000611253	0	0	0
2023/09/01-06090000614169	0	0	1
2023/09/01- 06090000625685	0	0	0
2023/09/01- 06090000735287	0	0	1
2023/09/01- 06090000814840	0	0	1
2023/09/01- 06090000845788	0	0	0
2023/09/01- 06090000872274	1	0	0

2.2 Pengolahan Data

2.2.1 Data Cleaning

Pada tahap data cleaning merupakan proses pembersihan dari data yang akan dipakai untuk penghapusan data dengan membuang missing value, duplikasi data dan memeriksa inkonsistensi data dan memperbaiki kesalahan pada data[27]. Proses pembersihan data dilakukan secara manual dengan bantuan software spreadsheet[28]. Pada data ini dilakukan pembersihan data seperti untuk bagian yang kosong (missing), data yang inkonsisten (salah input), serta data produk yang tidak memiliki kategori, seperti yang digambarkan pada Tabel 3.

Tabel 3. Data Cleaning

Keycode Trans	Cust nm	Prod nm	Sale Qty	cat	Remarks
2023/12/21-06090004878528		TELUR AYAM NEGERI CURAH /KG	2.030	Meat	Blank
2023/12/21-06090004878528		LEMON IMPORT /KG	162	Fruits	Blank
2023/12/21-06090004878528		JERUK MAND. POKKAM /KG	920	Fruits	Blank
2023/12/21-06090004878528		ROSE BRAND MINYAK GORENG PCH 2 LT	30	Bulk Product	Blank
2023/12/21-06090004878528		FIESTA CHICKEN NUGGET 500 GR	1	Dairy & Frozen	Blank
2023/09/15-06090003531626	31730542067 70008	SUNLIGHT CREAM LIME REF 350GR	6	Bulk Product	Inkonsisten
2023/09/15-06090003531626	31730542067 70008	BABY HAPPY S 10'S	1	Bulk Product	Inkonsisten
2023/09/15-06090003531626	A MUAMAR KADAFI	CHARM EXTDRYDAY WING 18'S	1	#N/A	Tidak memiliki kategori
2023/09/06-06018000010682	A MUAMAR KADAFI	GURAME MATI /KG	966	#N/A	Tidak memiliki kategori

2.2.2 Data Selection

Salah satu contoh *selection data* yang digambarkan pada Tabel 4 adalah menunjukkan item yang dibeli. *Noisy data* diatasi dengan memperbaiki data tersebut agar sesuai dengan data transaksi yang tertera pada data transaksi. Data yang telah melalui proses *data cleaning*, akan menjadi data yang siap diolah yang disimpan dalam file bertipe .csv atau .arff.

Tabel 4. Data Selection

Customer	Sale Day	Prod cd	Prod nm	Sale Qty
06008000106297	2024/02/20	1043581000	EXCLUSIVE HOTEL SOAP 500X15 G/CTN	1
06008000221214	2024/02/03	1090468000	COCA COLA KALENG 330 ML	24
06008000314648	2024/02/25	1083154000	NYU CREAM HAIR COLOR NATURAL BLACK	6
06018000004732	2024/02/07	1089155000	CLAS MILD 16'S-BKS	600
06090000402481		1097926000	ER MEN POLO SHIRT	3
06090000494586	2024/02/08	1060771000	CHOICE L GAS 230	56



Customer	Sale Day	Prod cd	Prod nm	Sale Qty
06090004963061	2024/02/17	1088129000	GR/BOTOL INDOMIE GORENG SPECIAL 85GR	40
06090004981816	2024/02/16	1099310000	PANASONIC LED TH-32G307G DIGITAL	1
06090004996480	2024/02/24	1099957000	SHARP FRIDGE 1D SJ-X187MG-DP/DB 166L	1

2.2.3 Data Transformation

Tahap Data Transformation merupakan proses mengubah format data awal menjadi sebuah format data standar untuk proses pembacaan data dengan algoritma pada program maupun tool yang digunakan. Berikut adalah hasil pengolahan data awal setelah melewati tahapan diatas untuk dijadikan dataset pada tahap selanjutnya, ditunjukkan pada Tabel 5.

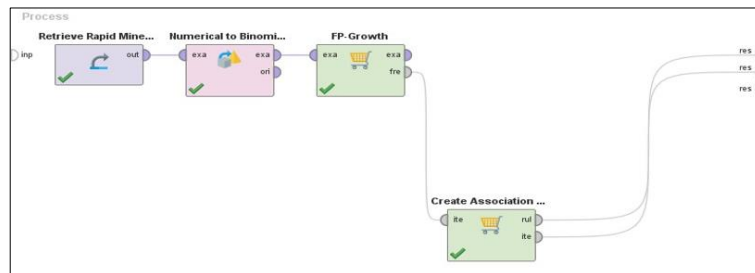
Tabel 5. Transformation

Sale Day	Prod nm
2024/02/20	EXCLUSIVE HOTEL SOAP 500X15 G/CTN
2024/02/03	COCA COLA KALENG 330 ML
2024/02/25	NYU CREAM HAIR COLOR NATURAL BLACK
2024/02/07	CLAS MILD 16'S-BKS
.....	ER MEN POLO SHIRT
2024/02/08	CHOICE L GAS 230 GR/BOTOL
2024/02/17	INDOMIE GORENG SPECIAL 85GR
2024/02/16	PANASONIC LED TH-32G307G DIGITAL
2024/02/24	SHARP FRIDGE 1D SJ-X187MG-DP/DB 166L
2024/02/28	RINNAI CAST IRON

2.3 Pemodelan

Pemodelan pada penelitian ini dilakukan dengan data mining teknik *Association rule* menggunakan algoritma Apriori. Teknik ini dipilih karena merupakan metode yang umum dipakai pada penelitian data mining untuk mencari seleksi aturan, maupun sebuah pola dari dataset besar.

Pada proses ini metode algoritma apriori diterapkan untuk menemukan keterkaitan itemset dengan keakurasian yang tepat, dengan menggunakan pengujian perhitungan dengan *tools* Rapid Miner.



Gambar 2. Proses Rapid Miner

3. HASIL DAN PEMBAHASAN

Pada bagian ini berisi hasil dan pembahasan dari topik penelitian, yang bisa di buat terlebih penerapan metode yang digunakan, baik secara sederhana dengan mengemukakan data yang ada pada penelitian. Bagian ini juga merepresentasikan penjelasan yang berupa penjelasan, gambar, tabel dan lainnya. Banyaknya kata pada bagian ini berkisar.

3.1 Analisa Data Penjualan

Tahapan awal yang dilakukan pada penelitian ini adalah mempersiapkan data, data yang akan diolah yaitu data penjualan produk retail. Sebagai sampel perhitungan tentang algoritma apriori dengan metode asosiasi penulis mengambil data yang akan diolah sebagai contoh yaitu sebanyak 500 data transaksi didapat dari 500 transaksi dengan total nominal penjualan tertinggi selama 6 bulan dalam rentang September 2023 – Februari 2024 seperti terlihat pada tabel 6.

Tabel 6 Data Transaksi Penjualan

Keycode Trans	Kategori	Sum of Sale amt
2023/09/01-06090000625685	Bulk Product	Rp 14.939.000



Keycode Trans	Kategori	Sum of Sale amt	
2023/09/02-06090003168550	Bulk Product	Rp	110.400.000
2023/09/02-06090003170200	Bulk Product	Rp	299.400.000
2023/09/03-06090003170200	Bulk Product	Rp	252.000.000
2023/09/04-06090000537686	Biscuit/Snacks	Rp	69.500
2023/09/04-06090000537686	CLEANING & LUGGAGE	Rp	4.500
2023/09/04-06090000537686	Dairy & Frozen	Rp	41.000
2023/09/04-06090000537686	Fruits	Rp	19.651
2023/09/04-06090000537686	Meat	Rp	15.762.000
2023/09/04-06090000537686	MILK/COFFEE/TEA	Rp	46.500
2023/09/04-06090000537686	Sauces&Spices	Rp	35.400
2023/09/05-06005000003439	HARDWARE	Rp	48.328.416
.....			
2024/02/27-06090000872274	MILK/COFFEE/TEA	Rp	2.698.900
2024/02/27-06090000872274	TEXTILE	Rp	5.286.718
2024/02/27-06090003128118	MILK/COFFEE/TEA	Rp	12.750.000
2024/02/27-06090005001193	DRINKS	Rp	48.750.000
2024/02/28-06090000576551	Bulk Product	Rp	5.212.500
2024/02/28-06090000576551	MILK/COFFEE/TEA	Rp	7.650.000
2024/02/28-06090001857762	Bulk Product	Rp	1.256.900
2024/02/28-06090001857762	Fruits	Rp	29.700
2024/02/28-06090001857762	MILK/COFFEE/TEA	Rp	15.720.000
2024/02/28-06090002136319	Bulk Product	Rp	53.440.000
2024/02/29-06090002080609	Bulk Product	Rp	20.986.800

Kemudian dari data pada Tabel 6 maka setelah dipisahkan didapat sebanyak 21 itemset antara lain seperti pada tabel 7.

Tabel 7. Itemset sampel

Itemset
APPLIANCE
Bakery
Biscuit/Snacks
Bulk Product
CLEANING & LUGGAGE
Dairy & Frozen
Detergent
DRINKS
Fish
FOOD RELATED
Fruits
H&B
HARDWARE
Meat
MILK/COFFEE/TEA
OFFICE & FURNITURE
PROFESSIONAL
MACHINERY
Ready To Eat
Sauces&Spices
TEXTILE
Vegetables

Dari proses mining yang dilakukan, didapatkan salah satu aturan asosiasi (association rule) berikut:
 {Milk, Coffee, Tea, Drinks, Detergent, Biscuit/Snacks} → Bulk Products
 Support: 12.5%
 Confidence: 78.6%
 Lift: 1.93

Artinya, dari total 5.972 transaksi yang dianalisis, sebanyak 12,5% transaksi mengandung kombinasi tersebut, dan dalam 78,6% dari kombinasi tersebut, konsumen juga membeli Bulk Products, yang menunjukkan hubungan asosiasi yang kuat.



Nilai lift sebesar 1.93 mengindikasikan bahwa peluang konsumen membeli Bulk Products meningkat hampir dua kali lipat ketika mereka membeli kombinasi produk dari lima kategori sebelumnya. Ini menunjukkan bahwa keterkaitan antarproduk bukan sekadar kebetulan, tetapi mengandung pola yang konsisten dalam kebiasaan belanja konsumen.

Temuan ini sangat relevan dalam konteks strategi penjualan ritel, terutama dalam perencanaan bundling produk, penempatan produk di rak (planogram), dan promosi terpadu. Retailer dapat memanfaatkan informasi ini untuk meningkatkan efisiensi penjualan dan kepuasan konsumen melalui pendekatan berbasis data (data-driven decision making).

3.2 Implementasi Algoritma Ariori

Sebagai perhitungan algoritma apriori dengan metode asosiasi penulis mengambil data yang akan diolah sebagai sampel perhitungan yaitu sebanyak 500 data. Penulis memberikan batasan nilai minimum support 0.1 atau sama dengan 10% dan confidence 0.5 atau sama dengan 50%. Iterasi satu dimulai dengan melakukan pembentukan kandidat 1 itemset (C1) dari 500 data transaksi tersebut dengan menghitung masing – masing jumlah supportnya. Menghitung nilai support untuk pola kombinasi 1 itemset yaitu dengan rumus sebagai berikut :

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

Dengan menggunakan rumus diatas, maka didapat perhitungan untuk masing – masing itemset sebagai berikut:

- a. Appliance = $\frac{7}{500} = 1,4\%$
- b. Bakery = $\frac{8}{500} = 1,6\%$
- c. Biscuit/Snack = $\frac{37}{500} = 7,4\%$
- d. Bulk Product = $\frac{473}{500} = 94,6\%$
- e. Cleaning & Luggage = $\frac{14}{500} = 2,8\%$
- f. Dairy & Frozen = $\frac{13}{500} = 2,6\%$
- g. Detergent = $\frac{34}{500} = 6,8\%$
- h. Drinks = $\frac{44}{500} = 8,8\%$
- i. Fish = $\frac{1}{500} = 0,2\%$
- j. Food Related = $\frac{6}{500} = 1,2\%$
- k. Fruits = $\frac{4}{500} = 0,8\%$
- l. H&D = $\frac{26}{500} = 5,2\%$
- m. Hardware = $\frac{7}{500} = 1,4\%$
- n. Meat = $\frac{3}{500} = 0,6\%$
- o. Milk Coffee Tea = $\frac{70}{500} = 14\%$
- p. Office & Furniture = $\frac{6}{500} = 1,2\%$
- q. Professional Machinery = $\frac{1}{500} = 0,2\%$
- r. Ready to Eat = $\frac{33}{500} = 6,6\%$
- s. Sauce & Spices = $\frac{25}{500} = 5\%$
- t. Textile = $\frac{8}{500} = 1,6\%$
- u. Vegetables = $\frac{0}{500} = 0\%$

3.3 Pembentukan Aturan Asosiasi

Kemudian selanjutnya dalam menentukan aturan asosiasi akan dihitung nilai confidence dari kombinasi tersebut, dengan ketentuan minimum nilai confidence sebesar 50%, menggunakan rumus berikut :

$$\text{Confidence P(A/B)} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \quad (2)$$

- a. Bulk Product, Milk/ Coffee/ Tea = $\frac{52473}{500} = 11\%$
- b. Milk/ Coffee/ Tea, Bulk Product = $\frac{5270}{500} = 74,3\%$

3.4 Pengujian Rapid Miner

Hasil pembentukan aturan Asosiasi dari pengujian Rapid Miner pada 500 data sample dapat dilihat pada gambar berikut:

AssociationRules

Association Rules

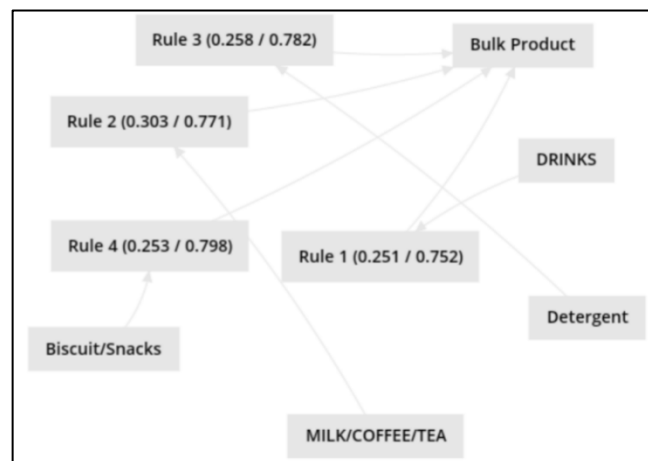
[MILK/COFFEE/TEA] --> [Bulk Product] (confidence: 0.743)

Gambar 3. Aturan Asosiasi hasil Rapid Miner

3.5 Perhitungan Menggunakan Dataset Uji

Dataset uji yang penulis gunakan diproses langsung menggunakan *tools* Rapid Miner, jika sebelumnya pada sampel perhitungan digunakan dataset sebanyak 500 data transaksi, maka untuk pengujian dataset yang digunakan adalah sebanyak 5972 data transaksi yang dimana data ini didapat adalah data yang memiliki total nominal penjualan yang diatas nilai rata-rata total nominal penjualan dari seluruh data transaksi pada periode September 2023 sampai dengan Februari 2024. Untuk jumlah itemset masih tetap yaitu sebanyak 21 itemset.

Pada data uji ini penulis menentukan nilai minimum support sebesar $\geq 25\%$, dan minimum confidence sebesar $\geq 65\%$. Hasil aturan asosiasi dari data uji yang diproses pada Rapid Miner menghasilkan grafik aturan seperti pada gambar 4.



Gambar 4. Asosiasi dari Dataset Uji pada Rapid Miner

Untuk hasil pembentukan aturan Asosiasi dari dataset uji sebanyak 5972 data pada Rapid Miner dapat dilihat seperti Gambar 5.

AssociationRules

Association Rules

[DRINKS] --> [Bulk Product] (confidence: 0.752)

[MILK/COFFEE/TEA] --> [Bulk Product] (confidence: 0.771)

[Detergent] --> [Bulk Product] (confidence: 0.782)

[Biscuit/Snacks] --> [Bulk Product] (confidence: 0.798)

Gambar 5 Hasil pembentukan aturan Asosiasi dari Dataset Uji pada Rapid Miner

Adapun untuk memudahkan pembacaan maka dibuatkan tabel seperti pada tabel 8.

Tabel 8. Aturan Asosiasi pada Data Uji

Itemset	Nilai <i>Support</i>	Nilai <i>Confidence</i>
Milk/ Coffee/ Tea, Bulk Product	30,3%	77,1%
Drinks, Bulk Product	25,1%	75,2%

Setelah melakukan percobaan mencari aturan Asosiasi pada dataset transaksi penjualan, algoritma apriori yang digunakan menghasilkan suatu aturan asosiasi dari setiap tahap yang dijalankan. Pada dataset sample ini ditentukan nilai minimum support sebesar $\geq 10\%$, dari data sampel perhitungan sebanyak 500 data transaksi penjualan, pada tahap iterasi pertama k-1 itemset didapatkan kumpulan sebanyak 2 itemset yang memenuhi syarat nilai support, yakni (Bulk Product) dengan nilai support 94,6% dan (Milk,Coffe,Tea) dengan nilai support sebesar 14%.

Kemudian pada kombinasi itemset pada tahap iterasi kedua, dari perhitungan pada data sampel hanya didapatkan 1 kombinasi k-2 itemset antara (Bulk Product) dengan (Milk, Coffe, Tea) yang memiliki nilai support sebesar 10,4%, dan selanjutnya yang sudah memenuhi syarat nilai minimum support diterjemahkan dalam aturan asoisasi yang terbentuk dengan menggunakan rule IF X Then Y, dan terdapat 2 buah aturan awal, kemudian dilakukan perhitungan nilai



confidence masing – masing aturan, dengan penentuan minimum confidence yaitu $\geq 50\%$. Hanya ada 1 rule yang memenuhi syarat untuk dijadikan sebagai aturan final dari hasil analisis dataset transaksi penjualan yang dipakai.

Hasil pengujian dengan data menggunakan tools Rapid Miner juga selaras dengan hasil perhitungan data sampel sebanyak 500 data transaksi, dengan menghasilkan nilai dan pembentukan aturan asosiasi yang sama.

Sementara jika menggunakan dataset uji yakni sebanyak 5972 data transaksi penjualan, dengan menaikkan nilai minimum support yakni sebesar $\geq 25\%$, pada tahap iterasi pertama k-1 itemset didapatkan untuk kumpulan itemset yang terbentuk yaitu (Bulk Product) dengan nilai support 74,1%, kemudian (Milk,Coffe,Tea) dengan nilai support 39,3%, kemudian (Drinks) dengan nilai support 33,4%, lalu ada (Detergent) dengan nilai support 33,1%, kemudian (Biscuit/Snacks) dengan nilai support 31,7%, dan (Sauces & Spices) dengan nilai support 27,5%. Disini terlihat perbedaan ketika diproses dengan data uji yang jumlahnya lebih banyak, maka dihasilkan itemset yang juga lebih banyak yang bisa memenuhi nilai minimum support yang ditentukan sebelumnya. Hal ini juga pada akhirnya berbanding lurus ketika diproses pada tools Rapid Miner dan terjadi peningkatan jumlah aturan Asosiasi yang terbentuk yaitu sebanyak 4 rules dengan varian itemset yang lebih banyak.

Pada hasil dari 500 data sampel yang dilakukan perhitungan manual, didapatkan kemungkinan aturan asosiasi diantaranya dimana ketika terjadi penjualan tinggi untuk produk (Milk,Coffe,Tea) pada satu waktu tertentu, maka ada kemungkinan disaat yang bersamaan juga terjadi kondisi yang sama untuk produk (Bulk Product), sedangkan jika menggunakan 5972 data transaksi pada dataset uji, aturan asosiasi juga berdampak pada kemungkinan terjadinya kombinasi itemset penjualan antara (Drinks) dengan (Bulk Product), kemudian (Detergent) dan (Bulk Product) serta (Biscuit/ Snacks) dan (Bulk Product).

4. KESIMPULAN

Penggunaan metode Asosiasi untuk analisa data transaksi penjualan melalui tahapan yang telah ditentukan dapat diterapkan dengan baik, dan mampu menghasilkan suatu aturan asosiasi baru dari dataset transaksi penjualan yang dapat dijadikan acuan dalam menentukan persediaan barang yang sesuai dengan kebutuhan pelanggan. Berdasarkan hasil penerapan association rule dengan menggunakan algoritma apriori pada data transaksi penjualan retail, pola penjualan produk dengan status permintaan tinggi dapat dilihat dari kombinasi penjualan produk yang terjual, pola asosiasi yang terbentuk yaitu dimana ketika terjadi penjualan tinggi untuk produk dalam kategori (Milk,Coffee,Tea), (Drinks), (Detergent), dan (Biscuit/Snacks) pada satu waktu tertentu, maka ada kemungkinan disaat yang bersamaan juga terjadi kondisi yang sama untuk produk yang termasuk dalam kategori (Bulk Products). Pengujian menggunakan tools Rapid Miner juga dapat menghasilkan wawasan yang serupa, dimana baik terhadap dataset sampel sebanyak 500 transaksi penjualan, maupun dengan dataset uji sebanyak 5972 transaksi penjualan, hasilnya linear dan berbanding lurus. Semakin banyak dataset uji yang digunakan, memungkinkan untuk mendapatkan kombinasi aturan asosiasi yang semakin variatif sehingga berpeluang memiliki referensi yang valid untuk diterapkan pada kebijakan marketing. Dalam penelitian ini, perlu dikembangkan dengan memperbanyak data yang akan digunakan dalam penelitian selanjutnya agar lebih banyak lagi kombinasi itemset yang terbentuk, pola pembelian konsumen dari penelitian ini menggunakan asosiasi dapat dimanfaatkan oleh perusahaan untuk meningkatkan strategi penjualan dengan mengombinasikan produk yang sering terjual bersamaan untuk membuat promosi ataupun paket belanja terhadap produk retail yang dijual. Penelitian selanjutnya disarankan untuk: 1) Menggunakan algoritma alternatif seperti FP-Growth atau Eclat yang memiliki efisiensi lebih tinggi dalam memproses dataset besar, guna membandingkan hasil dan performa dengan algoritma Apriori. 2) Mengembangkan sistem rekomendasi produk otomatis, yang memanfaatkan aturan asosiasi yang dihasilkan untuk meningkatkan pengalaman pelanggan dan penjualan dalam sistem e-commerce atau point of sale (POS). 3) Menerapkan teknik visualisasi data, seperti network graph atau heatmap, untuk menyajikan pola asosiasi yang ditemukan agar lebih mudah dipahami oleh pengguna non-teknis dalam pengambilan keputusan bisnis. 4) Mengintegrasikan data temporal, misalnya dengan mempertimbangkan waktu pembelian (hari, musim, atau periode promosi), untuk mengetahui dinamika pola pembelian konsumen secara lebih mendalam. 5) Melibatkan data multivariat lain seperti harga, stok, atau lokasi toko untuk memperluas konteks analisis dan menemukan insight yang lebih kompleks dalam perilaku konsumen. Dengan pengembangan di atas, diharapkan hasil penelitian di masa mendatang dapat memberikan kontribusi yang lebih luas, baik secara teknis maupun strategis, bagi dunia usaha retail dan pengembangan sistem intelijen bisnis.

REFERENCES

- [1] Miftahul Arifin, Fauzi Helmi, and D. F. Alamsyah, "Analisis Pola Asosiasi Penjualan Produk Ritel dengan Platform Google Colab," *JUSTIFY : Jurnal Sistem Informasi Ibrahimy*, vol. 3, no. 1, pp. 74–85, Aug. 2024, doi: 10.35316/justify.v3i1.5569.
- [2] S. Aulia Miranda, F. Fahrullah, and D. Kurniawan, "Implementasi Association Rule Dalam Menganalisis Data Penjualan Sheshop dengan Menggunakan Algoritma Apriori," *METIK JURNAL*, vol. 6, no. 1, pp. 30–36, Jul. 2022, doi: 10.47002/metik.v6i1.342.
- [3] T. Hidayat, I. R. Munthe, and A. P. Juledi, "Analisis Data Penjualan Menggunakan Algoritma Apriori pada Analisis Kopi," *INFORMATIKA*, vol. 12, no. 3, pp. 443–452, Dec. 2024, doi: 10.36987/informatika.v12i3.6064.
- [4] Andi Diah Kuswanto, Achmad Rizqullah Blessar, Abdul Goni, Arya Nibras Nayottama Sidiki, Oke Rizki Abdullah Haryu, and Hafid Anhar Hamiki, "Penerapan Algoritma Apriori Dalam Analisis Keranjang Belanja Retail Di Wilayah Jawa Barat," *Saturnus : Jurnal Teknologi dan Sistem Informasi*, vol. 2, no. 3, pp. 139–150, Jul. 2024, doi: 10.61132/saturnus.v2i3.208.



- [5] H. Fathiya Zahra and A. Triayudi, "Implementasi Business Intelligence untuk Memprediksi Penjualan Ritel Pada PT. Chelatama Perkasa Menggunakan Regresi Linear," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 4806–4814, May 2025, doi: 10.36040/jati.v9i3.13427.
- [6] I. Shantilawati, S. Zebua, and R. Tarmizi, "Penggunaan Digital Marketing Dalam Meningkatkan Penjualan Bisnis Retail," *JMARI*, vol. 5, no. 1, pp. 30–37, Feb. 2024, doi: 10.33050/jmari.v5i1.3130.
- [7] Lidiawati and Muhammad Mufti Imam Suyanto, "Analisis Dampak Usaha Ritel Modern Terhadap Riteltradisional," *Jurnal Ekonomi dan Bisnis Indonesia*, vol. 5, no. 2, pp. 1–9, Dec. 2020, doi: 10.37673/jebi.v5i02.848.
- [8] S. Suradi, H. Hakim, R. N. Tribowo, and N. Mukhlisah, "Analisa Pengaruh Proses Shrinkage (Penyusutan) Pada Produk Teh Pucuk Harum PT. Tirta Fresindo Jaya (Mayora Group)," *ILTEK : Jurnal Teknologi*, vol. 20, no. 01, pp. 32–36, May 2025, doi: 10.47398/iltek.v20i01.226.
- [9] A. D. Anggono, "Prediksi Shrinkage untuk Menghindari Cacat Produk pada Plastic Injection," *Media Mesin: Majalah Teknik Mesin*, vol. 6, no. 2, May 2015, doi: 10.23917/mesin.v6i2.2895.
- [10] F. I. Aryaniti, T. B. Santoso, Silvia, and M. Khairunnisa, "Penguatan Kemampuan Optimasi Parameter Proses Mesin Injeksi Plastik Otomotif untuk Minimasi Shrinkage," *I-Com: Indonesian Community Journal*, vol. 3, no. 4, pp. 2012–2018, Dec. 2023, doi: 10.33379/icom.v3i4.3477.
- [11] S. Hasan, A. Aryadi, and A. Suhud, "Analisis Komprehensif Pengaruh Retailing Mix Terhadap Keputusan Pembelian Konsumen: Systematic Review dan Meta-Analisis," *Equilibrium : Jurnal Ilmiah Ekonomi, Manajemen dan Akuntansi*, vol. 13, no. 2, p. 574, Sep. 2024, doi: 10.35906/equili.v13i2.2119.
- [12] A. Nur Rahmi and Yosaphat Ananda Mikola, "Implementasi Algoritma Apriori untuk Menentukan Pola Pembelian pada Customer (Studi Kasus : Toko Bakoel Sembako)," *Information System Journal*, vol. 4, no. 1, pp. 14–19, May 2021, doi: 10.24076/infosjournal.2021v4i1.561.
- [13] M. F. Haryanti *et al.*, "Pengaruh Data Mining, Strategi Perusahaan, Terhadap Laporan Kinerja Perusahaan," *Jurnal Portofolio : Jurnal Manajemen dan Bisnis*, vol. 3, no. 1, pp. 71–90, Jan. 2024, doi: 10.70704/jpjm.v3i1.285.
- [14] A. Wijaya, A. Faqih, D. Solihudin, C. L. Rohmat, and S. Eka Permana, "Penerapan Association Rules Menggunakan Algoritma Apriori Untuk Identifikasi Pola Pembelian," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 6, pp. 3871–3878, Feb. 2024, doi: 10.36040/jati.v7i6.8270.
- [15] D. Wulandari, B. Irawan, and A. Bachtiar, "Analisis Algoritma Apriori Untuk Mengidentifikasi Pola Penjualan Di Toko Sembako XYZ," *Jurnal Ilmu Teknik dan Komputer*, vol. 8, no. 1, p. 64, Jun. 2024, doi: 10.22441/jitkom.v8i1.009.
- [16] Y. Astuti and H. Novitasari, "Algoritma Apriori sebagai Penentu Pola Penjualan Produk Jeans," *Jurnal Ilmiah Edutic : Pendidikan dan Informatika*, vol. 9, no. 1, pp. 20–28, Nov. 2022, doi: 10.21107/edutic.v9i1.7416.
- [17] M. G. Saparudin and S. Sholihin, "Penggunaan Data Mining untuk Analisis Pola Pembelian Pelanggan Menggunakan Metode Association Rule Algoritma Apriori (Studi Kasus di Toko Waspada)," *Jurnal Teknologi Sistem Informasi dan Aplikasi*, vol. 6, no. 1, pp. 27–33, Jan. 2023, doi: 10.32493/jtsi.v6i1.26927.
- [18] R. N. Dianti and J. Zeniarja, "Implementasi Algoritma Apriori untuk Analisis Pola Pembelian Konsumen pada Toserba Yusuf Semarang," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 2, pp. 1013–1021, Jun. 2024, doi: 10.29100/jupi.v9i2.5421.
- [19] S. D. Rahmawati, A. B. Oktavia, F. S. A. Putri, and D. L. Fithri, "Penerapan Algoritma Apriori Untuk Menemukan Pola Asosiasi Pada Data Penjualan Retail Fashion," *SIMKOM*, vol. 10, no. 2, pp. 289–295, Jul. 2025, doi: 10.51717/simkom.v10i2.910.
- [20] S. Muharni and S. Andriyanto, "Penentuan Pola Penjualan Menggunakan Algoritma Apriori," *Digital Transformation Technology*, vol. 4, no. 1, pp. 60–71, Mar. 2024, doi: 10.47709/digitech.v4i1.3679.
- [21] I. Nawangsih and P. Purnamasari, "Analisis Pola Pembelian Produk Kecantikan Menggunakan Algoritma Apriori," *Jurnal Teknologi Informatika dan Komputer*, vol. 9, no. 1, pp. 537–546, Mar. 2023, doi: 10.37012/jtik.v9i1.1614.
- [22] F. C. Arifiyani and H. Pramaditya, "Peningkatan efektivitas pemasaran pada usaha retail melalui digitalisasi katalog dengan microsite," *Journal of Information System and Application Development*, vol. 1, no. 1, pp. 19–28, Mar. 2023, doi: 10.26905/jisad.v1i1.9860.
- [23] A. W. Bong and M. Mardiani, "Penerapan Data Analitik untuk Menentukan Pola Asosiasi Penjualan dengan Algoritma Apriori," *MDP Student Conference*, vol. 4, no. 1, pp. 31–38, Apr. 2025, doi: 10.35957/mdp-sc.v4i1.10930.
- [24] H. Sutrisno and N. A. S. Winarsih, "Klasifikasi Kategori Produk untuk Manajemen Keuangan Remaja menggunakan Algoritma Long Short-Term Memory," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 2, pp. 685–693, Dec. 2024, doi: 10.29408/edumatic.v8i2.27959.
- [25] N. A. Hibernastiar, A. F. Setiawan, and E. H. Susanto, "Penerapan Algoritma Apriori dalam Menentukan Rekomendasi Paket Produk," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, pp. 321–331, Jan. 2025, doi: 10.57152/malcom.v5i1.1782.
- [26] S. D. Rahmawati, A. B. Oktavia, F. S. A. Putri, and D. L. Fithri, "Penerapan Algoritma Apriori Untuk Menemukan Pola Asosiasi Pada Data Penjualan Retail Fashion," *SIMKOM*, vol. 10, no. 2, pp. 289–295, Jul. 2025, doi: 10.51717/simkom.v10i2.910.
- [27] A. muliawan Nur, M. Saiful2, H. Bahtiar, and Muhammad Taufik Hidayat, "Penerapan Algoritma K-Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasi," *Infotek: Jurnal Informatika dan Teknologi*, vol. 7, no. 2, pp. 478–488, Jul. 2024, doi: 10.29408/jit.v7i2.26283.
- [28] Y. Yunita, M. Fahmi, and S. Salmon, "Penerapan Algoritma K-Means Data Mining Pada Clustering Kelayakan Penerima UKT Dengan Normalisasi Data Model Z-Score," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1977–1986, Dec. 2024, doi: 10.47065/bits.v6i3.6475.