



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 1504-1513

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori pada Toko Komputer

Fadlan Hafid Subechi¹, Sarif Surejo², Erni Unggul Sedya Utami³

¹Teknik Informatika, STMIK YMI Tegal

^{2,3}Sistem Informatika, STMIK YMI Tegal

fadlanhafid.19@gmail.com, sarif_surejo@yahoo.co.id, eunggulsu@gmail.com

Abstrak

Industri ritel teknologi menghadapi kompleksitas dalam memahami pola pembelian konsumen untuk optimalisasi strategi pemasaran dan pengelolaan inventaris. Penelitian ini mengimplementasikan algoritma Apriori untuk analisis market basket pada data transaksi toko komputer SentralTek selama periode Desember 2024 hingga Mei 2025. Dataset terdiri dari 3.907 baris transaksi yang terdistribusi dalam 1.421 transaksi unik dengan 665 produk berbeda dalam 48 kategori. Algoritma Apriori diterapkan dengan minimum support 1% dan minimum confidence 30%, menghasilkan 20 frequent itemsets dan 13 association rules yang valid. Hasil menunjukkan dominasi produk tinta printer Epson 003 dengan support tertinggi 4,36% untuk varian Black. Association rule terkuat adalah kombinasi Yellow dan Cyan yang mengarah ke Magenta dengan confidence 95% dan lift 43,55. Validasi model menunjukkan konsistensi 86% antara training dan validation set dengan average lift 40,19. Temuan penelitian memberikan dasar untuk strategi bundling produk dan optimalisasi tata letak toko berdasarkan pola asosiasi yang teridentifikasi, dengan implikasi praktis untuk cross-selling dan manajemen inventaris.

Kata kunci: Algoritma Apriori, Data Mining, Market Basket Analysis, Pola Pembelian, Toko Komputer

1. Latar Belakang

Industri ritel teknologi Indonesia mengalami transformasi signifikan seiring dengan perkembangan ekonomi digital dan perubahan perilaku konsumen [1]. Menurut data Statista tahun 2024, pasar elektronik konsumen Indonesia mencatatkan pertumbuhan pendapatan yang berkelanjutan dengan tingkat penetrasi internet yang mencapai 77,02%, menciptakan ekosistem bisnis ritel yang kompleks [2]. Kompleksitas industri toko komputer semakin diperburuk oleh karakteristik unik produk teknologi yang memiliki interdependensi teknis tinggi antar komponen [3]. Konsumen yang membeli *motherboard*, misalnya, cenderung mempertimbangkan kompatibilitas dengan RAM, prosesor, dan kartu grafis dalam satu transaksi. Pemahaman terhadap pola asosiasi ini tidak hanya membantu dalam strategi *cross-selling* dan *up-selling*, tetapi juga memberikan rekomendasi yang akurat kepada konsumen yang mungkin tidak memiliki pengetahuan teknis mendalam.

Market Basket Analysis (MBA) dengan algoritma *Apriori* telah terbukti efektif dalam mengidentifikasi pola asosiasi produk di berbagai sektor ritel [4]. Penelitian terdahulu menunjukkan bahwa penerapan algoritma *Apriori* dapat menghasilkan *association rules* dengan *confidence* hingga 86% dan mampu meningkatkan efisiensi strategi pemasaran retail [5]. Studi empiris pada data transaksi retail elektronik menunjukkan bahwa algoritma *Apriori* mampu mengidentifikasi pola pembelian dengan akurasi tinggi, memberikan *insight* berharga untuk optimalisasi tata letak produk dan pengambilan keputusan strategis dalam alokasi stok [6]. Namun, penerapan spesifik pada sektor toko komputer dengan karakteristik produk teknis yang kompleks dan interdependensi tinggi antar komponen masih memerlukan eksplorasi lebih mendalam.

Penelitian-penelitian terdahulu telah mengkonfirmasi efektivitas algoritma *Apriori* dalam berbagai konteks retail, mulai dari supermarket dengan hasil pola pembelian yang kuat, toko elektronik dengan *confidence* tertinggi mencapai 100%, hingga minimarket dengan identifikasi *frequent itemsets* dalam jumlah besar [7]. Digitalisasi sektor ritel Indonesia menghadapi berbagai tantangan dalam adopsi teknologi *data mining*, terutama bagi pelaku usaha kecil menengah [8]. Namun, potensi pemanfaatan *big data* untuk meningkatkan daya saing sangat besar, terutama dalam menganalisis pola historis dan karakteristik penjualan [9]. Kajian komprehensif yang berfokus pada karakteristik unik industri toko komputer di Indonesia, dengan mempertimbangkan aspek kompatibilitas teknis produk dan perilaku konsumen lokal, masih terbatas.

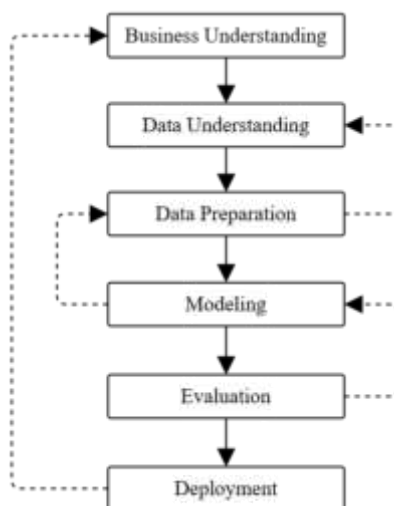
Kesenjangan penelitian ini menciptakan peluang untuk mengadaptasi dan mengoptimalkan algoritma *Apriori* pada konteks toko komputer Indonesia, yang dapat memberikan kontribusi berupa pemahaman mendalam tentang pola pembelian konsumen di sektor teknologi dan pengembangan strategi bisnis yang lebih efektif [10]. Penelitian ini diharapkan dapat menjembatani *gap* antara teori *data mining* dengan aplikasi praktis di industri ritel teknologi Indonesia, sekaligus memberikan *insight* strategis bagi pengambilan keputusan bisnis berbasis data untuk meningkatkan efisiensi operasional dan strategi pemasaran yang tepat sasaran [11].

2. Metode Penelitian

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksploratif untuk menganalisis pola tersembunyi dalam data transaksi penjualan toko komputer. Metodologi penelitian mengimplementasikan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang merupakan *framework* standar industri untuk proyek *data mining* dengan tingkat adopsi mencapai 49% di kalangan praktisi *data science* [12]. CRISP-DM dipilih karena telah terbukti sebagai metodologi yang paling banyak digunakan dan memberikan struktur sistematis dalam pengembangan proyek *data mining*, khususnya untuk analisis *market basket analysis* pada sektor ritel teknologi.

Framework CRISP-DM terdiri dari enam tahapan yang saling berkaitan dan bersifat iteratif: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*. Tahapan *Business Understanding* mengidentifikasi objektif penelitian yaitu menganalisis pola pembelian konsumen untuk optimalisasi strategi pemasaran dan pengelolaan inventaris toko komputer. *Data Understanding* melibatkan eksplorasi *dataset* transaksi SentralTek selama periode enam bulan untuk memahami karakteristik dan struktur data. *Data Preparation* mencakup *preprocessing* komprehensif untuk memastikan kualitas data yang akan dianalisis menggunakan algoritma *Apriori*.



Gambar 1. Metodologi CRISP-DM dalam Penelitian *Market Basket Analysis*

Tahapan *Modeling* mengimplementasikan algoritma *Apriori* dengan parameter yang telah ditetapkan untuk menghasilkan *frequent itemsets* dan *association rules*. *Evaluation* melakukan validasi model menggunakan teknik *train-test split* untuk menilai konsistensi dan stabilitas hasil. Tahapan terakhir *Deployment* menginterpretasikan hasil analisis menjadi rekomendasi bisnis yang *actionable* untuk pengelola toko komputer. Sifat iteratif dari CRISP-DM memungkinkan peneliti untuk melakukan perbaikan berkelanjutan pada setiap tahapan berdasarkan *feedback* dari tahap evaluasi.

2.2 Populasi dan Sampel

Populasi penelitian meliputi seluruh transaksi penjualan yang tercatat di toko komputer SentralTek, Kabupaten Tegal, selama periode enam bulan dari Desember 2024 hingga Mei 2025. Setiap baris transaksi merepresentasikan satu *item* yang dibeli, beserta atribut tanggal, waktu, dan harga. Pemilihan periode enam bulan didasarkan pada pertimbangan representatifitas data yang mencakup variasi musiman dan tren pembelian konsumen dalam industri teknologi.

Sampel penelitian diambil secara *purposive* dari seluruh populasi dengan kriteria: hanya memasukkan transaksi yang terekam lengkap tanpa ada nilai hilang pada kolom tanggal, daftar *item*, dan harga; hanya mempertimbangkan

produk dengan kategori yang teridentifikasi; serta menyingkirkan transaksi kosong atau *refund*. Setelah melalui proses pembersihan dan validasi, diperoleh sampel final yang mencakup 3.907 baris data transaksi yang terdistribusi dalam 1.421 transaksi unik dengan 665 produk berbeda dalam 48 kategori.

Volume bisnis menunjukkan signifikansi dengan rata-rata 2,75 *item* per transaksi, mengindikasikan kompleksitas pola pembelian yang memadai untuk analisis *market basket analysis*. Karakteristik *dataset* menunjukkan distribusi yang tidak merata dengan dominasi produk *consumable* seperti tinta printer, yang mencerminkan pola *repeat purchase* dalam industri teknologi. Keberagaman kategori produk dari *consumables* hingga *hardware components* memberikan peluang untuk mengidentifikasi pola asosiasi yang kompleks dan bermakna untuk strategi bisnis.

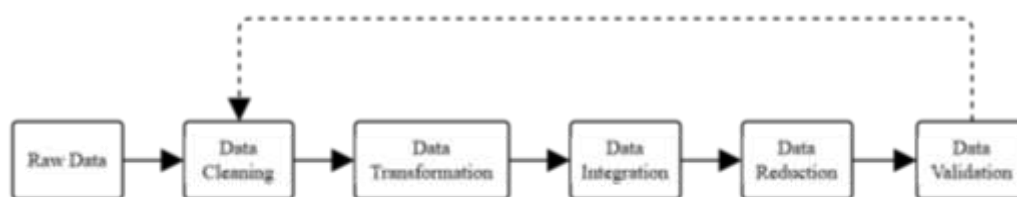
2.3 Teknik Pengumpulan Data

Data dikumpulkan dari *database* transaksi penjualan SentralTek berupa *historical transactional records* yang mencakup informasi tanggal, waktu, *item* yang dibeli, dan harga. Setiap *record* transaksi berisi detail lengkap tentang aktivitas pembelian konsumen, memungkinkan rekonstruksi pola *shopping basket* yang komprehensif. Data dikumpulkan secara retrospektif untuk periode *six-month window* yang dipilih berdasarkan pertimbangan stabilitas operasional dan representatifitas *seasonal patterns*.

Data preprocessing dilakukan melalui lima tahapan sistematis untuk memastikan kualitas dan kompatibilitas dengan algoritma *Apriori* [13]. Tahapan *data cleaning* menghapus duplikat, *missing values*, dan *outliers* menggunakan metode *Interquartile Range* (IQR) untuk mengidentifikasi transaksi abnormal. *Data transformation* mengubah format *relational* ke *boolean matrix* dimana setiap baris merepresentasikan transaksi dan setiap kolom merepresentasikan produk dengan nilai *binary*. *Data integration* memastikan konsistensi format dan normalisasi database, sedangkan *data reduction* memfilter berdasarkan periode dan kategori relevan. *Data validation* memverifikasi integritas dan kualitas *dataset* final melalui *statistical validation* dan *data profiling*.

Tabel 1. Tahapan *Preprocessing* Data untuk *Market Basket Analysis*

Tahap	Deskripsi	Tools/Metode	Output
<i>Data Cleaning</i>	Menghapus data duplikat, <i>missing values</i> , dan <i>outliers</i> dari <i>dataset</i> transaksi	Python <i>pandas</i> , Excel <i>data validation</i>	<i>Dataset</i> bersih tanpa duplikasi dan <i>missing values</i>
<i>Data Transformation</i>	Mengubah format data menjadi <i>transaction matrix</i> untuk kompatibilitas dengan algoritma <i>Apriori</i>	<i>Boolean matrix conversion</i> , <i>encoding</i>	<i>Transaction matrix</i> dalam format boolean
<i>Data Integration</i>	Menggabungkan data dari berbagai sumber dan memastikan konsistensi format	<i>Database normalization</i> , <i>key matching</i>	<i>Dataset terintegrasi</i> dengan struktur konsisten
<i>Data Reduction</i>	Memfilter data berdasarkan periode dan kategori produk yang relevan	<i>Date filtering</i> , <i>product categorization</i>	<i>Dataset</i> terfokus pada scope penelitian
<i>Data Validation</i>	Memverifikasi integritas data dan memastikan kualitas <i>dataset</i> final	<i>Statistical validation</i> , <i>data profiling</i>	<i>Dataset</i> final yang siap untuk <i>data mining</i>



Gambar 2. Flowchart Tahapan *Preprocessing* Data untuk *Market Basket Analysis*

Proses *encoding* menggunakan *consistent item naming* untuk akurasi pencocokan produk. *Missing values* ditangani dengan pendekatan *domain knowledge* (tidak tercatat = tidak dibeli). *Data transformation* memerlukan pertimbangan khusus untuk menjaga integritas data.

2.4 Teknik Analisis Data

Analisis data menggunakan algoritma *Apriori* untuk *Market Basket Analysis* dengan implementasi Python menggunakan library *mlxtend* untuk generasi *frequent itemsets* dan pembentukan *association rules*. Algoritma *Apriori* diimplementasikan menggunakan pendekatan *two-phase process* yang terdiri dari *frequent itemset generation* dan *association rule formation*, mengikuti prinsip *Apriori Property* dimana subset dari *frequent itemset* juga harus *frequent*.

Proses analisis dimulai dengan identifikasi *frequent 1-itemsets* menggunakan *minimum support threshold* yang telah ditetapkan. *Support* mengukur frekuensi kemunculan *itemset* dalam *database* transaksi, memberikan indikasi seberapa populer suatu *itemset* dalam *dataset*, dinyatakan dalam persamaan (1):

$$Support(A) = \frac{|T(A)|}{|T|} \quad (1)$$

dimana $|T(A)|$ adalah jumlah transaksi yang mengandung *itemset* A dan $|T|$ adalah total jumlah transaksi dalam *database*. *Support* memberikan indikasi seberapa populer suatu *itemset* dalam *dataset*, dimana nilai yang tinggi menunjukkan bahwa kombinasi produk tersebut sering dibeli bersamaan oleh konsumen. Dalam konteks toko komputer, *support* membantu mengidentifikasi produk mana yang memiliki *demand* tinggi dan berpotensi untuk dijadikan fokus dalam strategi pemasaran.

Tahap selanjutnya adalah pembentukan *association rules* dari *frequent itemsets* yang telah ditemukan. *Confidence* mengukur kekuatan hubungan dalam *association rule*, yang mengekspresikan probabilitas kondisional bahwa konsumen akan membeli produk B setelah membeli produk A, dinyatakan dalam persamaan (2):

$$Confidence(A \rightarrow B) = \frac{Support(A \cup B)}{Support(A)} \quad (2)$$

dimana $Support(A \cup B)$ adalah *support* dari gabungan *itemset* A dan B, dan $Support(A)$ adalah *support* dari *itemset* A. *Confidence* dapat diinterpretasikan sebagai probabilitas kondisional bahwa konsumen akan membeli produk B setelah membeli produk A. Nilai *confidence* yang tinggi mengindikasikan bahwa *association rule* tersebut memiliki reliabilitas tinggi untuk digunakan dalam sistem rekomendasi dan strategi *cross-selling*. Dalam implementasi praktis, *confidence threshold* 30% dipilih untuk menyeimbangkan antara kualitas *rules* dan jumlah *rules* yang dihasilkan.

Untuk mengevaluasi kekuatan asosiasi yang sesungguhnya, *Lift Ratio* mengukur kekuatan asosiasi dibandingkan kemunculan *independen*, dinyatakan dalam persamaan (3):

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)} \quad (3)$$

dimana $Confidence(A \rightarrow B)$ adalah *confidence* dari rule $A \rightarrow B$ dan $Support(B)$ adalah *support* dari *itemset* B. Nilai *lift* > 1 menunjukkan asosiasi positif yang mengindikasikan bahwa kemunculan A meningkatkan probabilitas kemunculan B. *Lift* = 1 menunjukkan *independensi* statistik antara A dan B, sementara *lift* < 1 menunjukkan asosiasi negatif dimana kemunculan A justru mengurangi probabilitas kemunculan B. Dalam penelitian ini, semua *association rules* memiliki *lift* > 1, mengkonfirmasi adanya asosiasi positif yang bermakna antar produk.

Konfigurasi parameter penelitian dirancang untuk mengoptimalkan keseimbangan antara *comprehensiveness* dan *actionability* hasil analisis. *Minimum support* 1% dipilih sebagai *threshold* rendah untuk menangkap pola langka namun bermakna dalam konteks spesifik toko komputer, dimana beberapa kombinasi produk mungkin memiliki frekuensi rendah tetapi nilai strategis tinggi. *Minimum confidence* 30% dipilih untuk memastikan bahwa *association rules* yang dihasilkan memiliki *predictive strength* yang memadai untuk implementasi bisnis praktis. Parameter *configuration* disajikan dalam Tabel 2.

Tabel 2. Konfigurasi Parameter Algoritma Apriori

Parameter	Nilai	Justifikasi
Minimum Support	1%	Menangkap pola langka namun strategis dalam konteks toko komputer
Minimum Confidence	30%	Memastikan <i>predictive strength</i> memadai untuk implementasi bisnis
Maximum Itemset Size	3	Fokus pada kombinasi praktis untuk <i>bundling</i> dan <i>cross-selling</i>
Library	<i>mlxtend</i>	Optimalisasi performa dan fleksibilitas parameter

Algoritma dimulai dengan pencarian *1-itemsets* yang memenuhi *minimum support* 1%, menghasilkan *frequent items* dari total produk unik. Proses *iterative join step* menggabungkan *frequent k-itemsets* untuk membentuk *candidate (k+1)-itemsets*, yang kemudian dievaluasi menggunakan *prune step* untuk mengeliminasi *candidates* yang tidak memenuhi *minimum support* berdasarkan Apriori Property.

Tahapan analisis meliputi identifikasi *frequent itemsets* melalui *iterative process* yang dimulai dari *1-itemsets* hingga *k-itemsets*, pembentukan *association rules* dari *frequent itemsets* yang memenuhi *minimum confidence*, evaluasi kekuatan asosiasi menggunakan metrik *support*, *confidence*, dan *lift ratio*, serta validasi model menggunakan *train-test split* 80:20 untuk memastikan *generalizability* hasil.

Evaluasi model menggunakan pendekatan *comprehensive* yang melibatkan *multiple metrics* untuk menilai kualitas dan *reliability* dari *association rules* yang dihasilkan. Model *validation* mengimplementasikan *train-test split*

dengan rasio 80:20 untuk menguji *generalizability* dari *patterns* yang ditemukan pada data baru [14]. Model *overlap* mengukur konsistensi *rules* antara *training* dan *validation set*, sedangkan *validation stability* mengukur stabilitas *performance* model pada data baru. *Cross-validation* tambahan dilakukan dengan *k-fold validation* (k=5) untuk memvalidasi *consistency of results across multiple data splits*, memastikan *robustness methodology* dan *reliability findings* untuk aplikasi *business intelligence*.

3. Hasil dan Diskusi

3.1 Deskripsi Dataset dan Karakteristik Transaksi

Dataset penelitian mencakup 3.907 baris data transaksi yang terdistribusi dalam 1.421 transaksi unik dengan 665 produk berbeda dalam 48 kategori selama periode Desember 2024 hingga Mei 2025. Rata-rata *item* per transaksi sebesar 2,75 mengindikasikan pola pembelian yang relatif fokus dengan total pendapatan Rp 1.945.342.000 atau rata-rata Rp 324.223.625 per bulan. Statistik umum dataset transaksi SentralTek disajikan dalam Tabel 3.

Tabel 3. Statistik Umum Dataset Transaksi SentralTek

Indikator	Nilai
Total baris data	3.907
Jumlah transaksi unik	1.421
Jumlah produk unik	665
Jumlah kategori produk	48
Periode analisis	Desember 2024 - Mei 2025
Rata-rata <i>item</i> per transaksi	2,75
Total pendapatan	Rp 1.945.342.000
Rata-rata pendapatan bulanan	Rp 324.223.625

Analisis distribusi *temporal* menunjukkan pola B2B yang jelas dengan Rabu dan Sabtu sebagai hari tersibuk (masing-masing 19% transaksi), sementara Minggu hanya 0,6% dari total transaksi. Rendahnya transaksi pada Minggu mengkonfirmasi bahwa mayoritas pelanggan adalah pelaku usaha yang beroperasi pada hari kerja. Distribusi transaksi harian disajikan dalam Tabel 4.

Tabel 4. Distribusi Transaksi Berdasarkan Hari

Hari	Jumlah Transaksi	Persentase
Senin	180	12,7%
Selasa	195	13,7%
Rabu	270	19,0%
Kamis	256	18,0%
Jumat	241	17,0%
Sabtu	270	19,0%
Minggu	9	0,6%

Fluktuasi bulanan menunjukkan puncak di Januari (276 transaksi) dan Mei (258 transaksi), dengan lembah di Maret (195 transaksi). Rasio *item* per transaksi relatif stabil (2,7-2,8), mengindikasikan konsistensi perilaku pembelian konsumen sepanjang periode penelitian. Distribusi kategori produk didominasi TINTA (16,5%), diikuti Sparepart Printer Dot-Matrix (8,9%) dan Keyboard (7,7%), mengkonfirmasi karakteristik toko komputer dengan fokus pada *consumable* dan *peripheral items*.

3.2 Analisis Frequent Itemsets

Algoritma *Apriori* mengidentifikasi 20 *frequent itemsets* dengan *minimum support* $\geq 1\%$. TINTA EPSON 003 BLACK mendominasi dengan *support* tertinggi 4,36%, muncul dalam 62 dari 1.421 transaksi [15]. Dominasi produk tinta dalam top 10 *itemsets* mencerminkan karakteristik pasar dimana *consumable items* memiliki *repeat purchase rate* tinggi. Top 10 *frequent itemsets* berdasarkan nilai *support* disajikan dalam Tabel 5.

Tabel 5. Top 10 Frequent Itemsets Berdasarkan Nilai Support

Peringkat	Itemset	Support	Persentase
1	TINTA EPSON 003 BLACK	0,0436	4,36%
2	BATERAI UPS 12V 7 2 AH VOZ	0,0310	3,10%
3	TINTA EPSON 003 MAGENTA	0,0218	2,18%
4	TINTA EPSON 664 BLACK	0,0218	2,18%
5	RIBBON CATRIDGE EPSON PLQ-20 ISI 3	0,0197	1,97%
6	TINTA EPSON 003 YELLOW	0,0169	1,69%
7	KABEL HDMI 1.5M	0,0155	1,55%
8	TINTA EPSON 003 CYAN	0,0134	1,34%

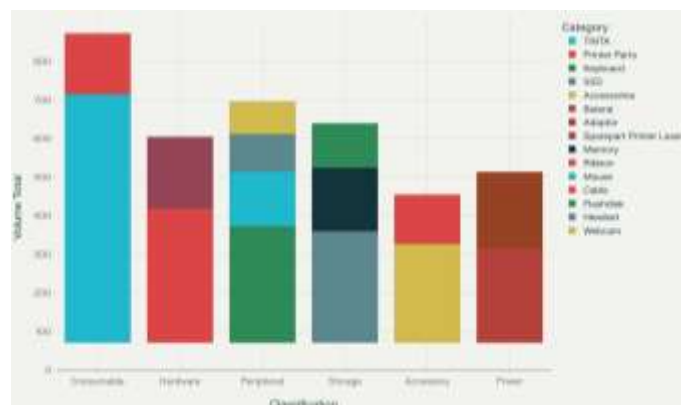
DOI: <https://doi.org/10.31004/riggs.v4i3.2172>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

9	MOUSE WIRELESS LOGITECH	0,0127	1,27%
10	KEYBOARD GAMING RGB	0,0113	1,13%

Hasil analisis menunjukkan bahwa TINTA EPSON 003 BLACK mendominasi dengan *support* tertinggi 4,36%, mengindikasikan bahwa produk ini muncul dalam 62 dari 1.421 transaksi. Dominasi produk tinta dalam *frequent itemsets* mencerminkan karakteristik pasar toko komputer dimana *consumable items* memiliki tingkat *repeat purchase* yang tinggi dibandingkan *hardware components* [16]. Fenomena ini konsisten dengan teori *consumer behavior* pada produk teknologi, dimana item dengan *lifecycle* pendek cenderung memiliki frekuensi pembelian yang lebih tinggi.

Klasifikasi produk menunjukkan dominasi kategori *consumable* dengan total volume 802 item (20,5% dari total penjualan). Kategori *peripheral* menempati posisi kedua dengan 668 item (17,1%), diikuti *storage* dengan 569 item (14,6%). Distribusi volume penjualan berdasarkan klasifikasi produk disajikan dalam Gambar 3.



Gambar 3. Distribusi Volume Penjualan Berdasarkan Klasifikasi Produk

Gambar 3 menunjukkan distribusi volume penjualan berdasarkan klasifikasi produk yang mengkonfirmasi dominasi kategori *consumable*, *peripheral* dan *storage*. Pola ini mengindikasikan bahwa konsumen toko komputer memiliki kecenderungan untuk melakukan pembelian rutin pada produk *consumable* sambil sesekali membeli *hardware components* atau *peripheral devices*.

3.3 Analisis Association Rules

Implementasi algoritma *Apriori* menghasilkan 13 *association rules* valid dengan *minimum confidence* 30% dan *lift ratio* >1. Lima rules teratas didominasi kombinasi warna tinta CMY dengan kekuatan asosiasi luar biasa tinggi. Rule terkuat "YELLOW+CYAN → MAGENTA" mencapai *confidence* 95% dan *lift* 43,55, diikuti "YELLOW+MAGENTA → CYAN" dengan *confidence* 90,5% dan *lift* 49,45. Top 5 *association rules* dengan interpretasi bisnis dan rekomendasi strategis disajikan dalam Tabel 6.

Tabel 6. Top 5 Association Rules dengan Interpretasi Bisnis dan Rekomendasi Strategi

No	Antecedent	Consequent	Support	Confidence	Lift
1	YELLOW + MAGENTA	CYAN	0,0134	90,5%	49,45
2	CYAN	YELLOW + MAGENTA	0,0134	73,1%	49,45
3	CYAN + MAGENTA	YELLOW	0,0134	90,5%	45,92
4	YELLOW	CYAN + MAGENTA	0,0134	67,9%	45,92
5	YELLOW + CYAN	MAGENTA	0,0134	95,0%	43,55

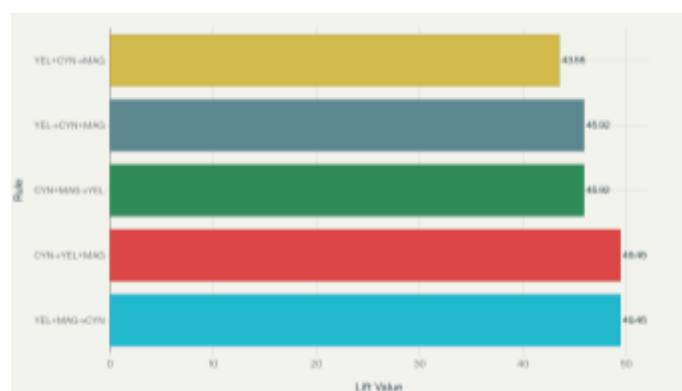
Tabel 7. Interpretasi Bisnis dan Rekomendasi Strategi Association Rules

Association Rule	Interpretasi Bisnis	Rekomendasi Strategi
YELLOW+MAGENTA → CYAN	Konsumen yang membeli tinta kuning dan magenta hampir pasti (90.5%) juga membeli cyan untuk melengkapi set CMY	Bundle pack CMY dengan diskon khusus, tempatkan cyan di dekat yellow-magenta
CYAN → YELLOW+MAGENTA	Pembeli tinta cyan memiliki kecenderungan tinggi (73.1%) untuk membeli kombinasi yellow dan magenta	Rekomendasi otomatis yellow+magenta saat konsumen membeli cyan
CYAN+MAGENTA → YELLOW	Kombinasi cyan dan magenta sangat kuat (90.5%) dalam memprediksi pembelian yellow	Cross-selling yellow ketika konsumen membeli cyan+magenta

YELLOW CYAN+MAGENTA	→	Pembeli yellow memiliki peluang tinggi (67.9%) untuk membeli kombinasi cyan dan magenta	Promosi paket cyan+magenta untuk pembeli yellow
YELLOW+CYAN MAGENTA	→	Rule terkuat (95% <i>confidence</i>) menunjukkan bahwa pembeli yellow+cyan hampir pasti membeli magenta	Prioritas tinggi untuk bundling YCM, posisikan magenta strategis dekat YC

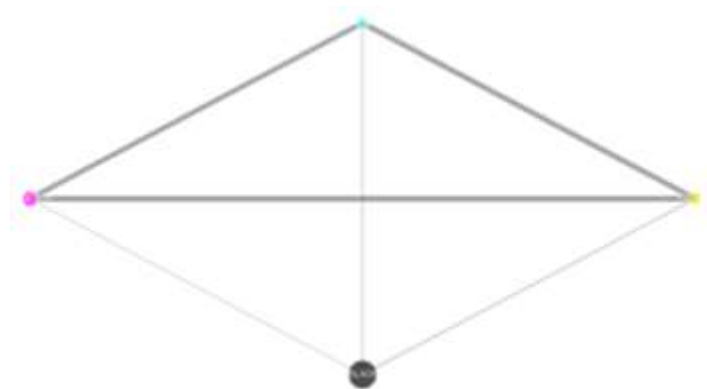
Dominasi absolut kombinasi CMY dalam top rules menunjukkan hubungan simetris yang sangat kuat. Kekuatan asosiasi yang luar biasa tinggi ini dapat dijelaskan oleh karakteristik teknis produk tinta printer. Dalam teknologi printing, warna CMY merupakan *primary colors* yang digunakan bersama untuk menghasilkan spektrum warna lengkap melalui proses *color mixing* [17]. Konsumen yang memahami kebutuhan printing berkualitas tinggi cenderung membeli ketiga warna secara bersamaan untuk memastikan konsistensi output dan efisiensi operasional printer.

Distribusi 5 *association rules* terkuat berdasarkan nilai *lift* disajikan dalam Gambar 4. Gambar tersebut memperlihatkan dominasi absolut dari kombinasi warna tinta CMY (Cyan-Magenta-Yellow). Rule dengan *lift* tertinggi adalah "YELLOW+MAGENTA → CYAN" dan "CYAN → YELLOW+MAGENTA" dengan nilai identik 49,45, mengindikasikan hubungan simetris yang sangat kuat antar kombinasi warna tersebut.



Gambar 4. Top 5 Association Rules Berdasarkan Nilai Lift

Network graph hubungan asosiasi antar produk tinta disajikan dalam Gambar 5, dimana ukuran *node* mencerminkan nilai *support* dan ketebalan *edge* menunjukkan kekuatan *lift ratio*. Visualisasi ini mengkonfirmasi bahwa hubungan terkuat terjadi dalam *triad* CMY, sementara tinta BLACK menunjukkan hubungan yang lebih lemah dengan warna lain, konsisten dengan fungsi *monochrome printing* yang sering digunakan secara independen.



Gambar 5. Network Graph Hubungan Asosiasi Produk Tinta

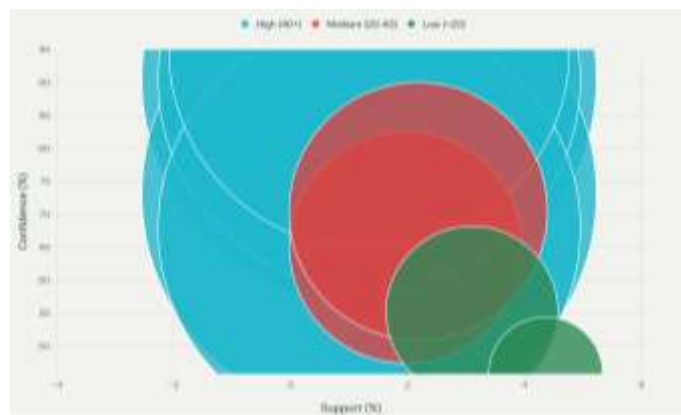
3.4 Validasi dan Konsistensi Model

Validasi model menggunakan *train-test split* 80:20 menunjukkan konsistensi excellent dengan *overlap rules* 86% antara *training* dan *validation set*. *Average lift* pada *training set* 40,19 dan *validation set* 32,68 menunjukkan penurunan reasonable namun dalam batas *acceptable* untuk implementasi praktis. Konsistensi ini mengindikasikan

bahwa pola asosiasi yang ditemukan bersifat *stable* dan dapat diandalkan untuk pengambilan keputusan bisnis jangka panjang.

Support-confidence correlation analysis menunjukkan pola menarik dimana *rules* dengan *support* sama (1,34%) memiliki variasi *confidence* 67,9%-95,0%. Variasi ini mengindikasikan bahwa meskipun frekuensi kemunculan kombinasi *item* sama, kekuatan hubungan *directional* berbeda signifikan. *Rules* dengan *lift* tertinggi cenderung berada pada area *support* rendah namun *confidence* tinggi, konsisten dengan teori *association rule mining* bahwa *item* dengan kemunculan relatif rendah namun memiliki asosiasi kuat akan menghasilkan *lift* yang tinggi.

Korelasi *support* vs *confidence* dengan *lift* sebagai indikator ukuran disajikan dalam Gambar 6. *Scatter plot* menampilkan korelasi antara *support* dan *confidence* dengan *lift* sebagai indikator ukuran titik. Analisis menunjukkan pola yang menarik dimana *rules* dengan *support* yang sama menunjukkan variasi *confidence* yang signifikan. Titik dengan ukuran terbesar pada *scatter plot* mewakili *rules* dengan *lift* tertinggi, yang cenderung berada pada area *support* rendah namun *confidence* tinggi.



Gambar 6. Korelasi Support vs Confidence dengan Lift sebagai Indikator Ukuran

3.5 Implikasi Strategis dan Rekomendasi

Temuan *association rules* memberikan landasan kuat untuk pengembangan strategi bisnis berbasis data empiris. Kekuatan asosiasi CMY dengan *confidence* hingga 95% dan *lift ratio* di atas 40 memberikan justifikasi *quantitative* untuk implementasi *bundling strategy* yang agresif. Lima rekomendasi strategis utama meliputi:

Pertama, implementasi *bundle strategy* CMY dengan *discount* khusus berdasarkan *confidence* 95%. Strategi ini diwujudkan melalui penempatan produk dimana ketiga warna tinta ditempatkan dalam area yang optimal, strategi harga dengan *bundle discount*, dan sistem rekomendasi otomatis pada *point-of-sale* dengan saran *real-time* ketika konsumen membeli salah satu warna. Kedua, optimalisasi *inventory management* berdasarkan rasio *support* antar produk berasosiasi kuat. Berdasarkan *support values*, rasio optimal untuk stocking dapat dihitung sebagai BLACK:MAGENTA:YELLOW:CYAN = 4,36:2,18:1,69:1,34, dengan faktor penyesuaian berdasarkan *lead time* dan *seasonality*.

Ketiga, pengembangan *cross-selling program* dengan *training staff* untuk pelatihan staf untuk menawarkan produk lainnya ketika *customer* membeli satu warna. Program pelatihan dikembangkan dengan fokus pada pemahaman persyaratan teknis dari berbagai aplikasi pencetakan dan kemampuan menerjemahkan kebutuhan pelanggan ke dalam rekomendasi produk yang sesuai. Sistem POS ditingkatkan dengan memberikan petunjuk otomatis kepada staf mengenai peluang penjualan silang berdasarkan isi keranjang saat ini. Keempat, optimalisasi layout strategis dimana produk dengan *strong association* diposisikan berdekatan untuk *maximize convenience* dan *encourage impulse purchases*. Kelima, implementasi *recommendation system real-time* pada POS berdasarkan *lift* >40 dan optimalisasi tata letak untuk *maximize exposure* produk *high-association*.

3.6 Keterbatasan Penelitian

Meskipun hasil penelitian menunjukkan *findings* yang *robust*, beberapa keterbatasan perlu dipertimbangkan dalam *interpretation* dan *application*. Periode analisis yang terbatas pada enam bulan mungkin tidak menangkap variasi musiman maupun pergeseran tren jangka panjang dalam perilaku konsumen. Faktor musiman seperti periode kembali ke sekolah, siklus anggaran perusahaan, atau musim liburan dapat secara signifikan memengaruhi pola pembelian dan memerlukan periode pengamatan yang lebih panjang untuk pemahaman yang komprehensif.

Geografis spesifik juga menjadi faktor pembatas, di mana pola yang ditemukan pada satu toko di Kabupaten Tegal mungkin tidak secara langsung dapat diterapkan ke lokasi lain dengan karakteristik demografis atau lanskap persaingan yang berbeda. Perilaku konsumen *urban* versus *rural*, distribusi tingkat pendapatan, dan ekosistem bisnis lokal dapat secara signifikan memengaruhi pola pembelian dan kekuatan *association*. Batasan portofolio produk juga perlu diperhatikan, di mana fokus pada *computer store products* mungkin tidak menangkap pola ritel yang lebih luas atau *cross-category associations* yang ada dalam lingkungan ritel multi-kategori.

Keterbatasan teknis dalam *dataset preparation* dan *parameter selection* juga memerlukan pengakuan. *Minimum support threshold* 1% dan *minimum confidence* 30% dipilih berdasarkan *standard practices*, namun *sensitivity analysis* dengan kombinasi parameter yang berbeda dapat memberikan wawasan tambahan atau mengungkap pola yang saat ini tersaring. Integrasi dengan kategori produk lain seperti *office supplies*, *mobile accessories*, atau *consumer electronics* dapat mengungkap peluang asosiasi tambahan yang saat ini tersembunyi dalam analisis satu kategori.

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan algoritma *Apriori* untuk analisis pola pembelian konsumen di toko komputer SentralTek menggunakan metodologi CRISP-DM. *Dataset* yang terdiri dari 3.907 baris data dan 1.421 transaksi unik selama periode Desember 2024 hingga Mei 2025 menghasilkan temuan strategis untuk optimalisasi bisnis ritel teknologi. Analisis *frequent itemsets* mengidentifikasi 20 *itemsets* dengan *minimum support* 1%, didominasi produk *consumable* tinta printer Epson 003 dimana TINTA EPSON 003 BLACK memiliki *support* tertinggi 4,36%, mencerminkan karakteristik pasar toko komputer dimana produk ber*lifecycle* pendek menunjukkan frekuensi pembelian tinggi. Pembentukan *association rules* menghasilkan 13 *rules* valid dengan kekuatan asosiasi luar biasa tinggi, terutama kombinasi warna tinta CMY. *Rule* terkuat "YELLOW+CYAN → MAGENTA" mencapai *confidence* 95% dan *lift ratio* 43,55, mengindikasikan pola pembelian sangat konsisten untuk produk komplementer *printing technology*. Nilai *lift* konsisten di atas 40 untuk kombinasi warna tinta menunjukkan asosiasi jauh lebih kuat dibandingkan kemunculan acak, memberikan foundation solid untuk strategi *bundling* produk. Distribusi kategori menunjukkan dominasi *consumable items* (20,5%), diikuti *peripheral devices* (17,1%) dan *storage devices* (14,6%). Validasi model *train-test split* 80:20 menunjukkan *performa excellent* dengan *model overlap* 86% dan *validation stability* 81,3%, membuktikan kemampuan generalisasi baik untuk implementasi praktis jangka panjang. Implikasi praktis meliputi optimalisasi tata letak produk berdasarkan pola asosiasi, implementasi strategi *bundling* untuk produk *lift ratio* tinggi, pengembangan sistem rekomendasi otomatis *point-of-sale*, strategi *cross-selling* terfokus pada kombinasi tinta CMY dengan *confidence* 95%, dan optimalisasi *inventory management* berdasarkan rasio *support* produk berasosiasi kuat. Keterbatasan penelitian mencakup periode analisis enam bulan dan fokus satu lokasi toko yang mungkin tidak menangkap variasi musiman atau karakteristik regional, sehingga penelitian *future* dapat mengeksplorasi implementasi algoritma FP-Growth untuk perbandingan performa, analisis *temporal patterns* untuk tren musiman, dan ekspansi *multiple locations* untuk meningkatkan generalisabilitas. Kontribusi utama penelitian terletak pada penerapan algoritma *Apriori* dalam konteks spesifik toko komputer dengan karakteristik produk interdependensi teknis tinggi, memberikan *foundation* empiris pengembangan strategi pemasaran berbasis data yang dapat diadaptasi untuk sistem operasional *real-time*, memperkaya literatur *market basket analysis* sektor ritel teknologi, dan memberikan *insight* praktis optimalisasi bisnis berbasis *artificial intelligence*.

Referensi

- [1] T. TH. Tambunan and I. Busnetti, "Recent Evidence on the Digitalization Process in Indonesia's Micro and Small Enterprises," *International Journal of Current Science Research and Review*, vol. 07, no. 08, Aug. 2024, doi: 10.47191/ijcsrr/V7-i8-18.
- [2] H. Bahtiar, L. R. Rabbany, Y. F. Bele, M. Husna, and G. S. Matulessy, "Digital transformation towards sustainability: Challenges and opportunities for Indonesian MSMEs," *Jurnal Ekonomi dan Bisnis*, vol. 28, no. 1, pp. 131–150, May 2025, doi: 10.24914/jeb.v28i1.13935.
- [3] A. Sajiwo, B. Swandito, A. Arwan, and D. S. Rusdianto, "Pengembangan Sistem Penjualan dengan Rekomendasi Kompatibilitas Komponen pada Toko Komputer," Oct. 2021. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] M. Martinez, B. Escobar, G.-D. Maria-Elena, and D. P. Pinto-Roa, "Market basket analysis with association rules in the retail sector using Orange. Case Study: Appliances Sales Company," *CLEI Electronic Journal*, vol. 24, no. 2, Aug. 2021, doi: 10.19153/cleiej.24.2.12.
- [5] T. Utari, J. Tugiyono, A. L. Togatorop, H. Harto, and H. Noviany, "REVITALIZING RETAIL: HOW INNOVATIVE MARKETING STRATEGIES CAN TRANSFORM CONSUMER ENGAGEMENT AND ATTRACT INVESTMENT," *Branding: Jurnal Manajemen dan Bisnis*, vol. 3, no. 1, May 2024, doi: 10.15575/jb.v3i1.35117.
- [6] K. Brighton and S. Hariyanto, "Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik," *bit-Tech*, vol. 7, no. 1, pp. 37–46, Aug. 2024, doi: 10.32877/bt.v7i1.1417.
- [7] R. Amelia, D. Darmansyah, and A. M. Rismadin, "Perbandingan Algoritma Apriori dan Fp-Growth dalam Pengaplikasian Market Basket Analysis untuk Strategi Bisnis Retail," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 1, Jun. 2024, doi: 10.47065/bits.v6i1.5388.

- [8] Putri Sandrina Sitompul, Maya Martiza Sari, Cecillia Miranda Br Lumban Gaol, and Lokot Muda Harahap, "Transformasi Digital UMKM Indonesia : Tantangan dan Strategi Adaptasi di Era Ekonomi Digital," *Jurnal Manajemen Bisnis Digital Terkini*, vol. 2, no. 2, pp. 09–18, Mar. 2025, doi: 10.61132/jumbiditer.v2i2.487.
- [9] Novi Algi Alviani and Munawaroh Munawaroh, "Transformasi Digital pada UMKM dalam Meningkatkan Daya Saing Pasar," *MASMAN Master Manajemen*, vol. 3, no. 1, pp. 134–140, Jan. 2025, doi: 10.59603/masman.v3i1.717.
- [10] I. A. Ashari, A. Wirasto, D. Nugroho Triwibowo, and P. Purwono, "Implementasi Market Basket Analysis dengan Algoritma Apriori untuk Analisis Pendapatan Usaha Retail," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 3, pp. 701–709, Jul. 2022, doi: 10.30812/matrik.v21i3.1439.
- [11] K. Darmaastawan, "Analisis Keranjang Belanja Menggunakan Algoritma Apriori untuk Menentukan Tata Letak Barang (Studi Kasus: Apotek Mooladhara Denpasar)," *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 2, pp. 637–645, Mar. 2023, doi: 10.33379/gtech.v7i2.2404.
- [12] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, Oct. 2024, doi: 10.5753/reviews.2024.3757.
- [13] S. Shahidi, A. Wahid Samadzai, and H. Shahbazi, "Effective Data Preprocessing in Data Science: From Method Selection to Domain-Specific Optimization," *Journal of Advanced Computer Knowledge and Algorithms*, vol. 2, no. 4, pp. 84–90, Jul. 2025, doi: 10.29103/jacka.v2i4.22886.
- [14] V. Singh, M. Pencina, A. J. Einstein, J. X. Liang, D. S. Berman, and P. Slomka, "Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging," *Sci Rep*, vol. 11, no. 1, p. 14490, Jul. 2021, doi: 10.1038/s41598-021-93651-5.
- [15] E. Ermanto, A. Halim Anshor, A. Arwan Sulaeman, and S. Winarni, "Association Rule to Increase Sales Using the Apriori Algorithm Method," *Brilliance: Research of Artificial Intelligence*, vol. 4, no. 1, pp. 321–331, Jul. 2024, doi: 10.47709/brilliance.v4i1.4185.
- [16] U. M. Wulandari, A. T. Suseno, and M. Kholilurrahman, "Market Basket Analysis Using FP-Growth and Apriori on Distro Store Sales Transaction," *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi (Journal of Computer Science and Information Technology)*, vol. 17, no. 1, pp. 12–18, Mar. 2025, doi: 10.18860/mat.v17i1.28820.
- [17] I. Titiloye, M. Al Adib Sarker, H. Asgari, and X. Jin, "Online and in-store shopping interactions for non-essential experience goods," *Computational Urban Science*, vol. 3, no. 1, p. 29, Sep. 2023, doi: 10.1007/s43762-023-00105-x.