

ANALISIS DATA MINING DALAM PENGELOLAAN PERSEDIAAN STOK DENGAN ALGORITMA RANDOM FOREST DAN APRIORI (STUDI KASUS: TOKO CERIA BABYSHOP)

Anzas Ibezato Zalukhu¹, Muhammad Iqbal², Darmeli Nasution³

Universitas Pembangunan Panca Budi, Medan

e-mail: ¹anzaszalukhu@gmail.com, ²muhammadiqbal@dosen.pancabudi.ac.id,

³darmelinasution@gmail.com

Abstract: *This research analyzes inventory management at Toko Ceria Babyshop by applying data mining techniques, specifically Random Forest and Apriori algorithms. Effective inventory management is crucial for aligning product availability with market demand, preventing overstocking or stockouts, and optimizing operational costs. Sales transaction data from June to December 2024, comprising 20,578 sales transactions, 3,593 purchase entries, 2,736 initial stock entries, and 1,331 final stock entries, were divided into 80:20 training and testing sets. The Random Forest implementation showed that weekly purchase quantity predictions were more effective than monthly predictions, evidenced by lower Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) values for weekly predictions (16.10, 1.76, 4.01) compared to monthly (39.68, 3.19, 6.30). Furthermore, the R-squared (R^2) value was higher for the weekly model (0.21) than the monthly (0.04), indicating better weekly prediction accuracy. The Apriori algorithm successfully identified product association patterns for both 2-itemsets and 3-itemsets, with all rules exhibiting lift values above 1, signifying positive relationships between products. This purchasing pattern information is highly beneficial for developing marketing strategies such as bundling, shelf arrangement, cross-selling promotions, and improved inventory planning.*

Keywords: *Data mining, Random Forest, Apriori, stok Inventory, Toko Ceria Babyshop*

Abstrak: Penelitian ini berfokus pada analisis pengelolaan persediaan stok di Toko Ceria Babyshop melalui penerapan teknik data mining menggunakan algoritma Random Forest dan Apriori. Efektivitas pengelolaan persediaan sangat krusial dalam bisnis untuk menyelaraskan ketersediaan produk dengan permintaan pasar, mencegah kelebihan atau kekurangan stok, dan mengoptimalkan biaya operasional. Data transaksi penjualan yang dikumpulkan dari Juni hingga Desember 2024 terdiri dari 20.578 transaksi penjualan, 3.593 entri pembelian, 2.736 entri stok awal, dan 1.331 entri stok akhir, yang kemudian dibagi menjadi set pelatihan dan pengujian dengan rasio 80:20. Hasil implementasi algoritma Random Forest menunjukkan prediksi kuantitas pembelian mingguan lebih efektif dibandingkan bulanan, ditunjukkan oleh nilai Mean Squared Error (MSE), Mean Absolute Error (MAE), dan Root Mean Squared Error (RMSE) yang lebih rendah pada prediksi mingguan (16.10, 1.76, 4.01) dibandingkan bulanan (39.68, 3.19, 6.30). Selain itu, nilai R-squared (R^2) juga lebih tinggi untuk model mingguan (0.21) dibandingkan bulanan (0.04), mengindikasikan akurasi prediksi mingguan yang lebih baik. Algoritma Apriori berhasil mengidentifikasi pola asosiasi produk, baik untuk 2-itemset maupun 3-itemset, dengan semua aturan memiliki nilai lift di atas 1, yang menunjukkan hubungan positif antar produk. Informasi mengenai pola pembelian ini sangat bermanfaat untuk pengembangan strategi pemasaran seperti bundling, penataan rak, promosi cross-selling, serta perencanaan persediaan stok yang lebih baik.

Kata kunci: Data mining, Random Forest, Apriori, Persediaan stok, Toko Ceria Babyshop

PENDAHULUAN

Pengelolaan persediaan yang efektif merupakan kunci keberhasilan dalam dunia bisnis yang kompetitif, karena memungkinkan perusahaan menyesuaikan ketersediaan produk dengan permintaan pasar, menghindari kekurangan atau kelebihan stok, serta mengoptimalkan biaya operasional. Dengan memanfaatkan teknologi seperti sistem inventori berbasis data dan pemantauan real-time, perusahaan dapat menganalisis tren pasar secara akurat dan merespons perubahan dengan cepat, sehingga meningkatkan efisiensi, kepuasan pelanggan, daya saing, serta menjaga profitabilitas dan keberlanjutan usaha (Harahap et al., 2023). Toko Ceria Babyshop adalah toko retail yang menjual berbagai perlengkapan bayi, namun menghadapi tantangan dalam pengelolaan stok seiring meningkatnya volume transaksi. Pengelolaan persediaan yang efisien perlu menjamin ketersediaan produk sesuai permintaan tanpa menimbulkan kelebihan stok yang membebani operasional. Untuk mengatasi hal ini, pendekatan proaktif dengan memanfaatkan teknik data mining dapat diterapkan, yaitu menggunakan algoritma Random Forest untuk memprediksi penjualan (forecasting stok) dan algoritma Apriori untuk menganalisis pola keterkaitan pembelian produk konsumen.

Data mining merupakan proses sistematis untuk menemukan informasi bernilai yang tersembunyi dalam kumpulan data besar (Putra & Wadisman, 2018; Ramadhani, 2025). Proses ini mengolah data mentah menjadi informasi bermakna dengan mengidentifikasi pola, tren, hubungan antar data, atau anomali yang sulit ditemukan secara manual (Sitorus et al., 2024; Sugito & Wahyuni, 2024; Suhardiansyah & Iqbal, 2025; Sutiono & Nasution, 2024; Wijaya, 2024). Data mining adalah metode umum untuk menggali informasi dari Big data (Nasution et al., 2022; Putera Utama Siahaan & Inda, 2024). Algoritma Random Forest adalah metode ensemble

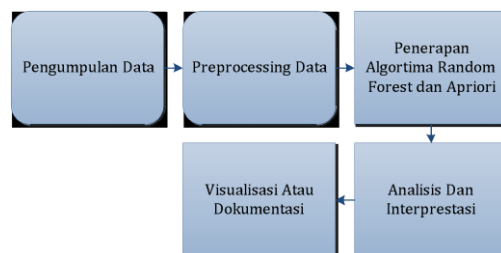
yang menggabungkan banyak Decision Tree dan menentukan hasil prediksi akhir melalui majority voting, sehingga lebih akurat dan mengatasi kelemahan dari penggunaan satu pohon keputusan (Sari et al., 2023; Syafira & Saepudin, 2023). Sedangkan Algoritma Apriori adalah metode association rule dalam data mining yang digunakan untuk menemukan pola atau keterkaitan antar item dalam data berskala besar, khususnya melalui analisis hubungan antar elemen dalam satu himpunan (Kusumo et al., 2019; Sunarti et al., 2021; Yuan, 2017).

Penelitian terkait menunjukkan efektivitas algoritma machine learning dalam berbagai domain. Penelitian Linawati et al., (2020) membandingkan algoritma Random Forest dan C4.5 untuk memprediksi prestasi akademik mahasiswa. Hasilnya, model klasifikasi Random Forest menunjukkan performa yang lebih unggul dengan akurasi 92,4%, presisi 91,4%, dan recall 92,4%. Selanjutnya, studi oleh Prasetya et al., (2022) menerapkan algoritma Apriori untuk menganalisis pola pembelian konsumen dari 84 transaksi dengan 119 atribut. Penelitian ini berhasil menemukan pola penjualan yang signifikan dengan nilai minimum support 0,25 (25%) dan confidence 0,25 (25%). Senada dengan itu, Budiman et al., (2024) juga memanfaatkan algoritma Apriori untuk menganalisis pola pembelian pada toko online Sayurklik. Temuan pola pembelian yang signifikan ini dapat menjadi dasar bagi pengembangan strategi pemasaran yang lebih efektif, seperti penawaran paket bundling dan rekomendasi produk, untuk meningkatkan layanan dan pemahaman terhadap perilaku konsumen. Integrasi algoritma Random Forest dan Apriori dalam pengelolaan persediaan memberikan pendekatan berbasis data yang kuat bagi ritel seperti Toko Ceria Babyshop. Random Forest memprediksi jumlah kebutuhan stok secara akurat berdasarkan data historis, sementara Apriori mengungkap pola keterkaitan antar produk dalam keranjang belanja.

Kombinasi keduanya memungkinkan toko menentukan jenis dan jumlah stok secara terpadu, memastikan ketersediaan produk yang sering dibeli bersamaan, mengisi stok sesuai pola belanja pelanggan, dan merespons perubahan permintaan dengan lebih cepat dan tepat.

METODE

Penelitian ini dilaksanakan di Toko Ceria Babyshop dengan periode pengumpulan data yang direncanakan berlangsung dari Juni hingga Desember 2024. Tujuan utama penelitian ini adalah menganalisis pengelolaan persediaan stok menggunakan metode data mining dengan algoritma *Random Forest* dan Apriori. Berikut tahapan-tahapan penelitiannya.



Gambar 1. Tahapan Penelitian

1. **Pengumpulan Data**
Pada tahap ini, data dikumpulkan sesuai kebutuhan penelitian. Data yang digunakan adalah data transaksi penjualan dari Juni hingga Desember 2024.
2. **Preprocessing Data**
Setelah data terkumpul, dilakukan tahap preprocessing untuk membersihkan data dari kesalahan atau informasi yang tidak relevan, seperti duplikasi, karakter yang hilang, atau elemen-elemen yang tidak dibutuhkan. Proses ini juga mencakup normalisasi data, seperti konversi ke huruf kecil, penghapusan tanda baca, serta formatisasi data menjadi bentuk yang siap untuk dianalisis.
3. **Implementasi Algoritma *Random Forest* dan Algoritma Apriori**

Tahap selanjutnya adalah penerapan algoritma Algoritma *Random Forest* dan Algoritma Apriori. Algoritma *Random Forest* dan Apriori merupakan metode yang efektif untuk menyelesaikan permasalahan berbasis data dengan pendekatan yang berbeda. *Random Forest* unggul dalam klasifikasi dan prediksi melalui pendekatan ensemble learning yang mampu menangani data kompleks dengan akurasi tinggi. Sementara itu, algoritma Apriori berfokus pada penemuan pola asosiatif dalam data, yang sangat bermanfaat untuk analisis hubungan antaritem, seperti dalam manajemen inventaris dan strategi pemasaran. Implementasi kedua algoritma ini dapat memberikan wawasan mendalam untuk mendukung pengambilan keputusan berbasis data yang lebih tepat dan efisien.

4. **Analisis dan Interpretasi Hasil**
Setelah implementasi algoritma selesai, tahap analisis dan interpretasi hasil klasifikasi dilakukan secara komprehensif. Peneliti mengevaluasi akurasi algoritma untuk memastikan performa dan validitasnya dalam mengolah data. Selain itu, tren atau pola signifikan yang teridentifikasi dalam data ulasan dianalisis guna memahami perilaku konsumen serta dinamika transaksi. Peneliti juga meninjau model data yang digunakan untuk memastikan relevansi dan interpretasi hubungan antarvariabel. Hasil dari analisis ini memberikan wawasan mendalam mengenai pola transaksi penjualan, yang selanjutnya dijadikan sebagai acuan strategis dalam memprediksi dan mengoptimalkan pengelolaan stok di masa depan, khususnya untuk kebutuhan operasional Toko Ceria Babyshop.
5. **Penyelesaian dan Dokumentasi**
Tahap akhir penelitian adalah penyelesaian proses analisis, di mana seluruh hasil yang diperoleh

didokumentasikan secara sistematis dan disusun dalam bentuk laporan penelitian. Laporan ini mencakup semua langkah penelitian, analisis data, dan interpretasi hasil secara komprehensif. Pada tahap ini, peneliti memastikan bahwa tujuan penelitian telah tercapai sesuai dengan rencana awal. Hasil analisis tidak hanya menjadi dasar untuk mengevaluasi strategi pemasaran dan meningkatkan kualitas layanan, tetapi juga memberikan wawasan yang bernilai dalam memprediksi kebutuhan pasar. Informasi ini dapat dimanfaatkan oleh pihak terkait, seperti pemilik usaha Toko Ceria Babyshop, untuk mengoptimalkan manajemen stok produk serta merespons permintaan konsumen dengan lebih efektif. Dengan demikian, penelitian ini memberikan kontribusi strategis yang signifikan bagi keberlanjutan dan pertumbuhan bisnis.

Teknik Analisis Data

Teknik analisis data yang digunakan meliputi:

1. Eksplorasi Data Analisis (EDA)
Eksplorasi Data Analisis (EDA) mencakup analisis deskriptif untuk memahami karakteristik data secara mendalam. Selain itu, visualisasi data digunakan untuk mengidentifikasi pola serta hubungan antar variabel, sehingga memberikan wawasan yang lebih jelas dalam pengambilan keputusan.
2. Pra-pemrosesan Data
Pra-pemrosesan Data mencakup beberapa langkah penting untuk memastikan kualitas data sebelum analisis lebih lanjut. Langkah pertama adalah pemilihan fitur yang relevan untuk prediksi, sehingga hanya variabel yang memiliki pengaruh signifikan yang digunakan. Selanjutnya, data dibagi menjadi set pelatihan dan pengujian untuk mengevaluasi kinerja model secara objektif. Terakhir, fitur numerik

diproses menggunakan Standard Scaler agar memiliki skala yang seragam, sehingga meningkatkan akurasi model.

3. Pemodelan

Dalam penelitian ini, model prediksi dibangun menggunakan dua algoritma utama, yaitu Random Forest dan Apriori.

Random Forest adalah *model ensemble* yang membangun banyak pohon keputusan dan mengambil rata-rata prediksi.

$$y = m = 1Mhm(x) \quad (1)$$

di mana:

M : adalah jumlah pohon.

Hm : adalah hasil prediksi dari masing-masing pohon

Algoritma Apriori menggunakan dua konsep utama *support* dan *confidence* (Hernawati, 2018; Madani et al., 2024).

Rumus untuk keduanya adalah sebagai berikut.

Rumus *1-itemset* =

Support A = Jumlah Transaksi Mengandung A / Total Transaksi x 100% (2)

Rumus *2-itemset* =

Support A,B = Jumlah Transaksi Mengandung A dan B / Total Transaksi x 100% (3)

Rumus *Convidence*

Convidence $P(B|A) = \text{Support A} \text{ dan } B / \text{Support A} \times 100\%$ (4)

4. Evaluasi Model

Metrik Evaluasi Menggunakan *MSE*, *RMSE*, *MAE*, dan *R-squared* untuk mengevaluasi performa model.

a. Mean Squared Error (MSE)

MSE mengukur rata-rata selisih kuadrat antara nilai aktual dan nilai prediksi. Semakin kecil nilai *MSE*, semakin baik modelnya.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

b. Root Mean Squared Error (RMSE)

RMSE adalah akar dari *MSE*, yang mengubah satuan kesalahan menjadi skala yang sama dengan data asli.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

c. *Mean Absolute Error (MAE)*

MAE mengukur rata-rata kesalahan absolut antara nilai aktual dan prediksi. Berbeda dengan *MSE* dan *RMSE*, *MAE* tidak mengkuadratkan selisih sehingga lebih robust terhadap outlier.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

d. *R-squared (R^2)*

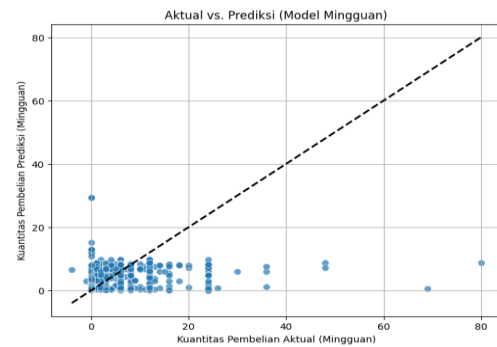
R^2 mengukur seberapa baik model menjelaskan variabilitas data. Nilai R^2 berkisar antara 0 hingga 1, di mana 1 berarti model sangat akurat, dan 0 berarti model tidak lebih baik dari rata-rata.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8)$$

HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini terdiri atas empat komponen utama, yaitu data penjualan, data pembelian, data stok awal, dan data stok akhir, yang mencakup rentang waktu dari bulan Juni hingga Desember 2024. Data penjualan mencatat sebanyak 20.578 transaksi, sedangkan data pembelian terdiri dari 3.593 entri. Adapun data stok awal mencakup 2.736 entri, dan data stok akhir mencakup 1.331 entri. Atribut-atribut yang terdapat dalam dataset meliputi kategori produk, ID produk, nomor faktur (*invoice*), tanggal penjualan, kuantitas penjualan, kuantitas pembelian, harga jual, tanggal pembelian, harga beli, jumlah pembelian, jumlah stok, serta nama produk. Dataset ini kemudian dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian dengan rasio pembagian sebesar 80:20.

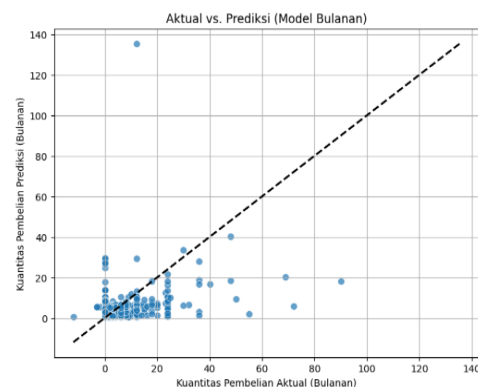
Algoritma *Random Forest*



Gambar 2 Prediksi Mingguan

IDProduk	NamaProduk	Rata2 QtyJual Historis Mingguan	PrediksiQtyBelMingguan
0	1	CESSA BABY BIRU	8.766667
1211	3103	KUMA KUMA BATH MITT	8.500000
2646	5313	FLUFFY STL KD PDK - PDK NB BLUE	8.333333
679	2017	FINI KAOS KAKI YB 367 / 152 / 405	7.750000
2648	5315	FLUFFY TOPI BULAT BLUE	8.000000
2107	4582	AMARA SPRAY COLOGNE FRESH 100G	0.000000
2109	4585	LINCO STL PDK - PDK S	0.000000
2113	4593	DLG (DGG1049) HIPSEAT TINI FREEFLOW	0.000000
2114	4594	DLG (DLK0286) KIJONG TARIK FOX	0.500000
2118	4598	EXPERT CARE DAILY MOISTURIZING SUN PRO SIG	0.000000

Gambar 3 Lanjutan Prediksi Mingguan



Gambar 4 Prediksi Bulanan

IDProduk	NamaProduk	Rata2 QtyJual Historis Bulanan	PrediksiQtyBelBulanan
0	1	CESSA BABY BIRU	37.571429
270	1344	MAMY LOVE BABY WIPES CHAMOMILE PINK SOS	18.000000
679	2017	FINI KAOS KAKI YB 367 / 152 / 405	17.714286
684	2023	MOCELL HAIR LOTION 100ML	17.714286
668	1996	SUN MARIE ROLL WCL 80G	18.428571
372	1549	BEEBIE NOURISHING BALM	18.000000
553	1796	MY BABY MINYAK TELON 150ML	18.571429
381	1573	NICE TISSUE 180G	25.571429
1967	44	LBBY SRG TGN & KAKI KARET	16.714286
57	1075	MY BABY MINYAK TELON 90ML	13.571429

Gambar 5 Lanjutan Prediksi Bulanan

Tabel 1 Hasil Evaluasi Model

Random Forest		
Metrik	Hasil Mingguan	Hasil Bulanan
Mean Squared Error (MSE)	16.10	39.68
Mean Absolute Error (MAE)	1.76	3.19
Root Mean Squared Error (RMSE)	4.01	6.30
R-squared: (R^2)	0.21	0.04

Berdasarkan hasil evaluasi model *Random Forest*, performa prediksi pada data mingguan lebih baik dibandingkan data bulanan. Hal ini terlihat dari nilai *mean squared error* (MSE) yang lebih rendah pada model mingguan (16.10) dibandingkan bulanan (39.68), begitu pula dengan *mean absolute error* (MAE) dan *root mean squared error* (RMSE) yang masing-masing lebih kecil pada data mingguan (1.76 dan 4.01) dibandingkan data bulanan (3.19 dan 6.30). Nilai *r-squared* (R^2) sebesar 0.21 untuk model mingguan menunjukkan bahwa model mampu menjelaskan sekitar 21% variasi data aktual, sedangkan hanya 4% pada model bulanan, yang menandakan prediksi bulanan jauh lebih lemah. Secara keseluruhan, model *Random Forest* lebih efektif digunakan untuk memprediksi kuantitas pembelian mingguan dibandingkan bulanan.

Algoritma Apriori

Aturan asosiasi dalam analisis ini menggunakan nilai *support* sebesar 0,05 dan *confidence* sebesar 0,6. Berdasarkan parameter tersebut, berikut adalah hasil aturan asosiasi yang dihasilkan dari data yang dianalisis.

Aturan Asosiasi 2-itemset

	Jika Membeli (Antecedent)	Maka Akan Membeli (Consequent)	support	confidence	lift
45	BESI SEGITIGA AYUNAN	PER AYUNAN TUNGGAL 3-9 KG	0.051087	0.607500	11.211538
44	PER AYUNAN TUNGGAL 3-9 KG	BESI SEGITIGA AYUNAN	0.051087	0.846154	11.211538
294	USAGI SRG TGN SBL	USAGI SRG TGN PLS	0.051087	0.607500	3.309535
213	LIBBY STL KD PJG - PJG NB 3-6	USAGI SRG TGN PLS	0.066038	0.608096	3.001011
269	MOM UUNG ASI BOOSTER 60 KAPSUL	SLEEK BABY BN CLEANSER 150ML BTL	0.051087	0.611111	2.272904

Gambar 6 Hasil Aturan Asosiasi 2-itemset

Hasil tersebut menunjukkan hasil analisis asosiasi produk menggunakan algoritma Apriori aturan 2-itemset, yang mengidentifikasi pasangan produk yang sering dibeli bersamaan. Misalnya, pelanggan yang membeli "PER AYUNAN TUNGGAL 3-9 KG" sangat mungkin juga membeli "BESI SEGITIGA AYUNAN" dengan *confidence* 84,6% dan *lift* 11,21, menunjukkan hubungan yang sangat kuat. Nilai *lift* di atas 1 menandakan asosiasi positif antar produk. Informasi ini dapat dimanfaatkan untuk strategi *bundling*, penataan rak, atau promosi *cross-selling*, serta untuk perencanaan persediaan stok agar produk-produk yang saling berkaitan tersedia bersamaan dan menghindari kekosongan stok saat permintaan meningkat.

Aturan Asosiasi 3-itemset

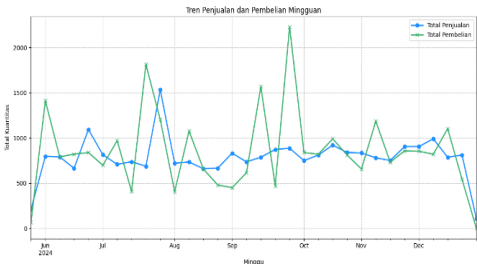
	Jika Membeli (Antecedents)	Maka Akan Membeli (Consequents)	support	confidence	lift
953	CESSA BABY BIRU, RIMOSHI CLN PJG BABYWEAR S	RIMOSHI BAJU PDK BABYWEAR NB	0.051321	0.619040	4.374003
952	CESSA BABY BIRU, RIMOSHI BAJU PDK BABYWEAR NB	RIMOSHI CLN PJG BABYWEAR S	0.051321	0.650000	4.175758
1336	RIMOSHI BAJU PDK BABYWEAR NB, MINYAK KAYU PUTIH	RIMOSHI CLN PJG BABYWEAR S	0.051321	0.619040	3.576912
724	MINYAK KAYU PUTIH AYAM 40ML, SOFTEX MATERNITY	BUMBOO KALDU AYAM 90G	0.051087	0.647059	2.867043
1217	MAKUKU DRY CARE MSH, MINYAK KAYU PUTIH AYAM 40ML	MOELL BODY LOTION 185GR	0.056604	0.857143	2.836206

Gambar 7 Aturan Asosiasi 3-itemset

Gambar 7 tersebut menampilkan hasil analisis asosiasi produk menggunakan algoritma Apriori dengan aturan 3-itemset, yang bertujuan untuk mengidentifikasi kombinasi tiga produk yang sering dibeli bersamaan. Contohnya, pelanggan yang membeli "CESSA BABY BIRU" dan "RIMOSHI CLN PJG BABYWEAR S" cenderung juga membeli "RIMOSHI BAJU PDK BABYWEAR NB", dengan *confidence* sebesar 61,9% dan *lift* 4,37, yang menunjukkan asosiasi kuat antar produk tersebut. Semua aturan dalam tabel memiliki nilai *lift* di atas 1, yang

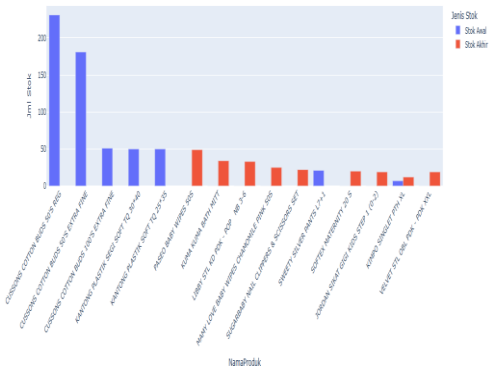
mengindikasikan hubungan positif. Informasi ini sangat berguna tidak hanya untuk strategi penjualan seperti *bundling* dan *cross-selling*, tetapi juga untuk pengelolaan persediaan, agar produk-produk terkait selalu tersedia secara bersamaan, mengurangi risiko kekosongan stok, dan meningkatkan efisiensi penyediaan barang di toko.

Visualisasi Tren dan Rekomendasi

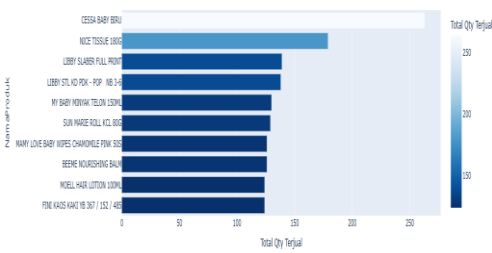


Gambar 8 Tren Penjualan, Pembelian, Dan Stok

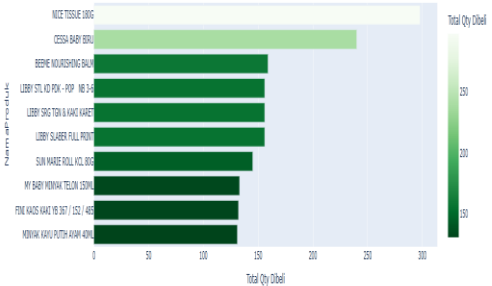
Hasil Grafik tersebut menunjukkan tren penjualan dan pembelian mingguan sepanjang Juni hingga Desember 2024. Terlihat bahwa jumlah pembelian (garis hijau) cenderung lebih fluktuatif dan memiliki lonjakan tinggi pada beberapa minggu, seperti awal Agustus dan akhir September. Sementara itu, penjualan (garis biru) lebih stabil dan konsisten setiap minggunya. Perbedaan ini menunjukkan bahwa pembelian seringkali tidak sebanding langsung dengan penjualan, sehingga penting bagi toko untuk menyesuaikan jumlah pembelian dengan tren penjualan agar stok tetap efisien dan tidak berlebih.



Gambar 9 Perbandingan Stok



Gambar 10 Produk Terlaris



Gambar 11 Produk terbanyak dibeli

	NamaProduk	Stok	Rata2PenjualanMingguan	MingguPasokan	StatusStok
414	VELVET STL OBL PAG - PUG S	12	1.44	8.3	Stok Berlebih (Lambat Bergerak)
412	VELVET STL OBL POK - PUG RIB XL	12	1.42	8.5	Stok Berlebih (Lambat Bergerak)
1233	SANDAL BABY BUSA PETAK	10	1.17	8.5	Stok Berlebih (Lambat Bergerak)
917	STERIMAR BABY NOSE HYGIENE 50ML	12	1.40	8.6	Stok Berlebih (Lambat Bergerak)
1311	LIBBY TOPI BULAT ROCCO GIRL	12	1.40	8.6	Stok Berlebih (Lambat Bergerak)

Gambar 12 Potensi Peningkatan Manajemen Stok Mingguan

Tabel tersebut mengindikasikan adanya lima produk dengan kondisi *overstock* dan tingkat perputaran yang rendah (*slow moving*), ditunjukkan oleh rata-rata penjualan mingguan yang rendah (1,17–1,44 unit) serta estimasi waktu habis stok mencapai 8,3 hingga 8,6 minggu. Situasi ini berpotensi menimbulkan inefisiensi dalam manajemen persediaan, seperti peningkatan biaya penyimpanan, risiko penurunan kualitas produk, serta akumulasi stok yang tidak sesuai dengan tingkat permintaan.

	NamaProduk	Stok	Rata2PenjualanBulanan	BulanPasokan	StatusStokBulanan
481	USAGI CUN POP POLOS	12	4.57	2.6	Stok Berlebih (>2.5 Bln)
869	DODO BABY SOOTHER 342	6	2.29	2.6	Stok Berlebih (>2.5 Bln)
993	SUGARBABY PREMIUM ROCKER 10 IN 1	9	3.50	2.6	Stok Berlebih (>2.5 Bln)
1268	THANK YOU BAG PREMIUM L	19	7.43	2.6	Stok Berlebih (>2.5 Bln)
365	MINYAK KAYU PUTIH LANG 15ML	8	3.00	2.7	Stok Berlebih (>2.5 Bln)

Gambar 13 Potensi Peningkatan Manajemen Stok Bulanan

	NamaProduk	Stok	Rata2PenjualanMingguan	MingguPasokan	StatusStok
3	ROYAL BABY WALKER 202	0	1.60	0.0	Stok Habis (Permintaan Ada)
4	ROYAL BABY WALKER 020 NEO	0	1.50	0.0	Stok Habis (Permintaan Ada)
5	ROYAL BABY WALKER 230 A NEO	0	1.33	0.0	Stok Habis (Permintaan Ada)
6	PMB BABY WALKER BW 703	0	1.14	0.0	Stok Habis (Permintaan Ada)
7	PMB BABY WALKER BW 506	3	1.55	1.9	Stok Rendah (Risiko Habis)

Gambar 14 Rekomendasi Stok Mingguan

	NamaProduk	Stok	Rata2PenjualanBulanan	BulanPasokan	StatusStokBulanan
3	ROYAL BABY WALKER 202	0	3.20	0.0	Stok Habis (Permintaan Ada - Bulanan)
4	ROYAL BABY WALKER 020 NEO	0	1.50	0.0	Stok Habis (Permintaan Ada - Bulanan)
5	ROYAL BABY WALKER 230 A NEO	0	2.29	0.0	Stok Habis (Permintaan Ada - Bulanan)
6	PMB BABY WALKER BW 703	0	1.33	0.0	Stok Habis (Permintaan Ada - Bulanan)
11	DUG (DLB3372) BANTAL PEANG BEARIE	0	1.00	0.0	Stok Habis (Permintaan Ada - Bulanan)

Gambar 15 Rekomendasi Stok Bulanan

SIMPULAN

Penelitian ini berhasil menerapkan algoritma Random Forest dan Apriori dalam pengelolaan persediaan stok pada Toko Ceria Babyshop, dengan hasil evaluasi menunjukkan bahwa model Random Forest lebih efektif untuk memprediksi kuantitas pembelian mingguan dibandingkan bulanan berdasarkan nilai MSE, MAE, RMSE, dan R-squared yang lebih baik pada data mingguan. Sementara itu, algoritma Apriori berhasil mengidentifikasi pola pembelian asosiatif produk, baik untuk 2-itemset maupun 3-itemset, dengan nilai lift di atas 1, mengindikasikan hubungan positif antar produk yang dapat dimanfaatkan untuk strategi penjualan seperti bundling dan cross-selling, serta optimalisasi ketersediaan stok. Analisis tren penjualan dan pembelian mingguan juga menunjukkan bahwa pembelian cenderung lebih fluktuatif daripada penjualan, menegaskan pentingnya penyesuaian jumlah pembelian dengan tren penjualan demi efisiensi stok. Meskipun model Random Forest menunjukkan hasil yang lebih baik pada data mingguan, nilai R-squared sebesar 21% masih menunjukkan keterbatasan dalam kemampuan model menjelaskan variasi data pembelian. Oleh karena itu, disarankan untuk menggunakan data

tambahan yang lebih komprehensif, seperti data faktor eksternal (promosi, musim, atau kondisi pasar), guna meningkatkan akurasi prediksi. Selain itu, eksplorasi penggunaan model lain seperti Gradient Boosting, XGBoost, atau model deret waktu seperti LSTM dapat dipertimbangkan untuk memperoleh hasil yang lebih optimal.

DAFTAR PUSTAKA

- A., Handoko, T. S., Salsabila, T. D., & Siregar, J. (2024). Analisis Pola Pembelian Pada Toko Online Sayurklik Berbasis Website Menggunakan Metode Apriori. *Journal of Sciencetech Research and Development*, 6(1), 705–714. <https://doi.org/10.56670/JSRD.V6I1.383>
- Harahap, F., Fahrozi, W., Adawiyah, R., Siregar, E. T., & Harahap, A. Y. N. (2023). Implementasi Data Mining dalam Memprediksi Produk AC Terlaris untuk Meningkatkan Penjualan Menggunakan Metode Naive Bayes. *JURNAL UNITEK*, 16(1), 41–51. <https://doi.org/10.52072/unitek.v16i1.541>
- Hernawati. (2018). Analisis Market Basket Dengan Algoritma Apriori. *Ikraith-Informatika*, 2(18), 13–17.
- Kusumo, H., Sedyono, E., & Marwata, M. (2019). Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi. *Walisongo Journal of Information Technology*, 1(1), 49. <https://doi.org/10.21580/wjit.2019.1.1.4000>
- Linawati, S., Nurdiani, S., Handayani, K., & Latifah, L. (2020). Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Random Forest Dan C4.5. *Jurnal Khatulistiwa Informatika*, 8(1). <https://doi.org/10.31294/JKI.V8I1.7827>
- Madani, M. I., Padmo, A., Masa, A., &

- Setyadi, H. J. (2024). Perbandingan Metode Apriori Dan Frequent Pattern. 12(2).
- Nasution, D., Sirait, D. N., & Wardani, I. (2022). OPTIMASI JUMLAH CLUSTER METODE K-MEDOIDS BERDASARKAN NILAI DBI PADA PENGELOMPOKAN DATA LUAS TANAMAN DAN PRODUKSI KELAPA SAWIT DI SUMATERA UTARA. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 09 No.02, 381. <https://klik.ulm.ac.id/index.php/klik/article/view/464>
- Prasetya, T., Yanti, J. E., Purnamasari, A. I., Dikananda, A. R., & Nurdianawan, O. (2022). Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori. *Informatics For Educators And Professional: Journal Of Informatics*, 6(1), 43. <https://doi.org/10.51211/itbi.v6i1.1688>
- Putra Utama Siahaan, A., & Inda, M. (2024). Analysis of Age and Gender Classification Using Decision Tree Model in the Context of Nursing Homes | *Journal of Information Technology, computer science and Electrical Engineering. Journal of Information Technology, Computer Science and Electrical Engineering (JITCSE)*, Vol. 1, No. 2, 149–152. <https://doi.org/10.61306/jitcse.v1i2>
- Putra, R. R. (Randi), & Wadisman, C. (Cendra). (2018). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *Intecoms*, 1(1), 72–77. <https://doi.org/10.31539/INTECOMS.V1I1.141>
- Ramadhani, S. E. (2025). Analisis Data Pengguna Perpustakaan Digital Menggunakan AI Data Mining. *Prosiding Sains Dan Teknologi*, 4(1), 373–380. <https://jurnal.pelitabangsa.ac.id/index.php/SAINTEK/article/view/5666>
- Sari, L., Romadloni, A., Listyaningrum, R., Studi, P., Informatika, T., Cilacap, P. N., & Soetomo, J. (2023). Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random Forest. *Infotekmesin*, 14(1), 155–162. <https://doi.org/10.35970/INFOTEK MESIN.V14I1.1751>
- Sitorus, Z., Syahputra, I., Angkat, C. I., & Sartika, D. (2024). Implementation of K-Means Clustering for Inventory Projection. *International Journal of Science, Technology & Management*, 5(3), 673–678. <https://doi.org/10.46729/IJSTM.V5I3.856>
- Sugito, B., & Wahyuni, S. (2024). Optimasi Strategi Penjualan Am2000 Tirtamart Dengan Algoritma Apriori Untuk Mengidentifikasi Produk Favorit Pelanggan. *Bulletin of Information Technology (BIT)*, 5(4), 278–286. <https://doi.org/10.47065/BIT.V5I4.1707>
- Suhardiansyah, & Iqbal, M. (2025). Optimization of Kaayana Store Inventory through Transaction Pattern Analysis Using the Apriori Algorithm. *Journal Of Data Science*, 3(01), 1–9. <https://doi.org/10.58471/JDS.V3I01.6398>
- Sunarti, S., Handayanna, F., & Irfiani, E. (2021). Analisa Pola Penjualan Makanan Dengan Penerapan Algoritma Apriori. *Techno.Com*, 20(4), 478–488. <https://doi.org/10.33633/tc.v20i4.4715>
- Sutiono, S., & Nasution, D. (2024). Analysis of Student Graduation at SMK Negeri 1 Stabat Using the C4.5 Algorithm. *Journal of Information Technology, Computer Science and Electrical Engineering*, 1(3), 160–163. <https://doi.org/10.61306/JITCSE.V1I3.96>
- Syafira, M. N., & Saepudin, D. (2023). Prediksi Return Saham dengan Metode Random Forest dan

-
- Penerapannya untuk Seleksi Portofolio. EProceedings of Engineering, 10(5). <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/21364>
- Wijaya, A. T. (2024). Penerapan Metode Clustering Dengan Algoritma K-Means Pada Sistem Pendeteksi Pencucian Uang Application Of Clustering Methods With K-Means Algoritma On Web-Based Banking Money Cheating. 3(September), 398–406.
- Yuan, X. (2017). An improved Apriori algorithm for mining association rules. AIP Conference Proceedings, 1820(September). <https://doi.org/10.1063/1.4977361>