

SURAT TUGAS
0743/B.01/LPPM-UBSI/V/2026

Tentang

PENELITIAN YANG DIPUBLIKASIKAN DALAM JURNAL ILMIAH
Periode Maret-Agustus 2026

Menulis pada Jurnal INSAN (Journal of Information Systems Management Innovation)
Volume 6 No. 1 Juni 2026 (e-ISSN : 2777-1385)

Judul :

Analisa Pola Asosiasi Produk Pada Transaksi Penjualan Retail Online Menggunakan Algoritma Apriori

- Menimbang :
1. Bahwa perlu diadakan pelaksanaan Tridharma Perguruan Tinggi dalam bentuk Penelitian.
 2. Untuk Keperluan pada butir 1 (satu) diatas, maka perlu dibentuk tugas yang berkaitan dengan penelitian yang dipublikasikan dalam Jurnal Ilmiah.

MEMUTUSKAN

- Pertama :
- Menugaskan kepada saudara
1. 199609156 Elly Mufida
 2. 200209764 Luci Kanti Rahayu
 3. 201709220 Doni Andriansyah
 4. 201209696 Eva Rahmawati
 5. 201103250 Hylenarti Hertyana

Sebagai Penulis yang mempublikasikan Penelitiannya pada Jurnal Ilmiah.

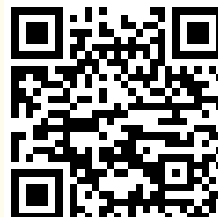
- Kedua :
- Mempunyai tugas sbb:
Melaksanakan Tugas yang diberikan dengan penuh rasa tanggung jawab.

- Ketiga :
- Keputusan ini berlaku sejak tanggal ditetapkan, dengan ketentuan apabila dikemudian hari terdapat kekeliruan akan diubah dan diperbaiki sebagaimana mestinya.

Jakarta, 4 Mei 2026

LPPM Universitas Bina Sarana Informatika

Ketua



Agus Junaidi, M. Kom

Tembusan

- Ka. LPPM Universitas Bina Sarana Informatika
- Arsip
- Ybs

Analisa Pola Asosiasi Produk Pada Transaksi Penjualan Retail Online Menggunakan Algoritma Apriori

Ana Nazmi Aulia¹, Elly Mufida², Luci Kanti Rahayu³, Eva Rachmawati⁴, Hylenarti Hertiyana⁵, Doni Andriansyah⁶

^{1,2,3,4,5,6}Universitas Bina Sarana Informatika

Jl. Kramat Raya No.98. Jakarta, Indonesia

e-mail: ¹ananazmiaulia@gmail.com, ²elly.elm@bsi.ac.id, ³luci.lkr@bsi.ac.id, ⁴eva.ehw@bsi.ac.id,
⁵hylenarti.hha@bsi.ac.id, ⁶doni.dad@bsi.ac.id

Artikel Info : Diterima : 06-05-2026 | Direvisi : 05-06-2026 | Disetujui : 08-06-2026

Abstrak - Pertumbuhan e-commerce telah menciptakan peluang signifikan bagi sektor ritel online untuk lebih memahami perilaku konsumen. Namun, banyak bisnis masih menghadapi tantangan dalam mengidentifikasi hubungan antar produk yang sering dibeli bersamaan. Studi ini bertujuan untuk mengungkap pola asosiasi produk dalam transaksi penjualan menggunakan algoritma Apriori. Dataset yang digunakan adalah dataset Fashion Retail Sales yang diperoleh dari Kaggle, terdiri dari 3.400 transaksi, dengan 166 transaksi yang digunakan setelah pra-pemrosesan. Metode yang diterapkan adalah penambangan data dengan teknik analisis keranjang belanja menggunakan RapidMiner. Parameter yang digunakan adalah dukungan minimum 0,07% dan kepercayaan minimum 0,5%. Hasil menunjukkan asosiasi signifikan antara Tanktop dan Celana Pendek, serta antara Celana Pendek dan Dompot. Pola-pola ini dapat berfungsi sebagai dasar untuk mengembangkan strategi rekomendasi produk, penggabungan produk, dan penempatan produk yang lebih efektif. Temuan ini penting karena dapat membantu bisnis dalam meningkatkan strategi pemasaran dan meningkatkan pengalaman belanja konsumen di platform ritel online.

Kata Kunci : Algoritma Apriori, Data Mining, E-commerce, Analisa Keranjang Belanja

Abstracts - The growth of e-commerce has created significant opportunities for the online retail sector to understand consumer behavior better. However, many businesses still face challenges in identifying relationships between products frequently purchased together. This study aims to reveal patterns of product associations in sales transactions using the Apriori algorithm. The dataset used is the Fashion Retail Sales dataset obtained from Kaggle, consisting of 3,400 transactions, with 166 transactions used after preprocessing. The applied method is data mining with the market basket analysis technique using RapidMiner. The parameters used are a minimum support of 0.07% and a minimum confidence of 0.5%. The results show significant associations between Tanktop and Shorts, as well as between Shorts and Wallet. These patterns can serve as a foundation for developing product recommendation strategies, bundling, and more effective product placement. The findings are important as they can assist businesses in improving marketing strategies and enhancing consumer shopping experiences on online retail platforms.

Keywords : Apriori Algorithm, Data Mining, E-commerce, Market Basket Analysis, Retail

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi pada era digital telah membawa transformasi yang signifikan terhadap berbagai sektor kehidupan, termasuk industri perdagangan ritel. Perubahan perilaku konsumen yang semakin beralih ke platform digital mendorong pelaku bisnis untuk mengadopsi strategi berbasis data dalam menghadapi persaingan yang semakin ketat. Menurut data Bank Indonesia, Indeks Penjualan Riil (IPR) nasional mencapai level 242,9 pada April 2023, mencatat rekor tertinggi sejak awal pandemi Covid-19. Perdagangan besar dan eceran memberikan kontribusi sebesar Rp2.702,4 triliun, setara dengan 12,94% dari Produk Domestik Bruto (PDB) Indonesia, yang menunjukkan kontribusi sangat besar dan signifikan terhadap



perekonomian nasional.(Fuadi et al., 2024). Dalam konteks ini, retail online telah menjadi salah satu pilar utama dalam ekosistem perdagangan modern. Kemudahan akses, jangkauan pasar yang luas, serta efisiensi transaksi menjadikan platform e-commerce sebagai pilihan utama bagi konsumen dalam memenuhi kebutuhan sehari-hari (Nurdiansyah et al., 2024). Pertumbuhan pesat retail online menghasilkan volume data transaksi yang sangat besar setiap harinya. Data transaksi penjualan yang terekam dalam sistem informasi toko online mengandung informasi berharga mengenai pola pembelian konsumen, preferensi produk, serta kebiasaan belanja yang dapat dimanfaatkan untuk pengambilan keputusan bisnis strategis (Brighton & Hariyanto, 2024). Masalah umum yang dihadapi oleh pengusaha di bidang ritel adalah bagaimana cara memenuhi keinginan pelanggan sambil mencapai keuntungan maksimal. Untuk mengatasi masalah ini, perusahaan memerlukan sumber informasi yang baik, salah satunya adalah dengan menganalisis data transaksi yang ada.(Fuadi et al., 2024).

Market Basket Analysis (MBA) merupakan teknik yang umum digunakan dalam menganalisis perilaku konsumen di sektor ritel. Metode ini menganalisis hubungan antar produk yang dibeli konsumen dalam satu transaksi dengan tujuan menemukan pola pembelian yang sering terjadi. Melalui MBA, bisnis dapat memahami pola pembelian konsumen secara lebih komprehensif dan merancang strategi pemasaran yang lebih tepat sasaran, seperti penawaran bundling atau rekomendasi produk yang relevan.(Brighton & Hariyanto, 2024)(Safitry et al., 2025). Dalam kerangka data mining, algoritma Apriori merupakan salah satu algoritma yang paling umum digunakan pada Market Basket Analysis karena kemampuannya dalam mengidentifikasi aturan asosiasi yang relevan dari dataset berukuran besar (Tarigan et al., 2022). Algoritma ini dikembangkan oleh Agrawal dan Srikant pada tahun 1994 dan bertujuan untuk menemukan kumpulan item yang sering muncul (frequent itemset) dalam pembentukan aturan asosiasi (Ananda et al., 2023). Sebagai bagian dari data mining, algoritma Apriori bekerja dengan membentuk kombinasi kandidat item dan menguji apakah kombinasi tersebut memenuhi parameter minimum yang ditetapkan, yaitu support dan confidence (Sutoyo et al., 2024). Support menggambarkan frekuensi kemunculan kombinasi item dalam dataset, sementara confidence menunjukkan tingkat kepercayaan atau kekuatan hubungan antar item dalam transaksi (Mubarak et al., 2019).

Berbagai penelitian terdahulu telah membuktikan efektivitas algoritma Apriori dalam menganalisis pola asosiasi produk pada sektor ritel. Penelitian yang dilakukan oleh Nur dkk. (2025) menerapkan algoritma Apriori dalam mengidentifikasi pola perilaku belanja konsumen di toko D&D Mart dan berhasil menemukan kombinasi produk kopi, deterjen, dan gula dengan nilai support 31% dan confidence 100%, yang menunjukkan pola pembelian produk yang sangat kuat terjadi secara bersamaan (Nur et al., 2025) Penelitian serupa dilakukan oleh Fuadi dkk. (2024) yang menganalisis pola pembelian konsumen di Toko Ritel DMART menggunakan algoritma Apriori berbasis website dengan 600 data transaksi dan menemukan 6 aturan asosiasi dengan nilai minimum support 6,6% dan confidence 65% (Fuadi et al., 2024). Selain itu, Safitry dkk. (2025) berhasil mengidentifikasi association rule yang kuat antar produk pada data transaksi e-commerce Toko Retail X dengan dataset 527.765 record data, menghasilkan kombinasi produk dengan confidence hingga 90,48% dan lift ratio 17,64 (Safitry et al., 2025).

Penelitian lain yang dilakukan oleh Jufri dkk. (2024) menerapkan algoritma Apriori untuk menentukan pola penjualan barang di Aciak Mart dengan 50 transaksi, menggunakan nilai minimum support 6% dan minimum confidence 15%, menghasilkan aturan asosiasi dengan confidence tertinggi 35,71% untuk pembelian Sampoerna Mild 16 dan Surya 16 (Jufri et al., 2024) Hanapi dkk. (2023) menerapkan algoritma Apriori untuk rekomendasi produk pada toko online Rizvenastore berbasis website, dengan 61 data transaksi dan minimum support 60% serta minimum confidence 50%, menghasilkan tiga rule asosiasi yang dapat diimplementasikan sebagai rekomendasi produk best-sellerb(Hanapi & Sari, 2023) Sementara itu, Shafa dan Perdana (2024) menerapkan algoritma Apriori pada data transaksi penjualan di Toko Berkah Berlimpah dengan 240 data transaksi, menggunakan minimum support 30% dan minimum confidence 60%, menghasilkan tiga aturan asosiasi yang dapat dimanfaatkan untuk pengambilan keputusan bisnis (Tam et al., 2024).

Industri fashion adalah kegiatan kreatif yang terkait dengan kreasi desain pakaian, desain alas kaki, dan desain aksesoris mode lainnya, produksi pakaian dan aksesorisnya, konsultasi lini produk fashion, serta distribusi produk fashion. Saat ini di Indonesia industri fashion termasuk dalam salah satu sub-sektor industri kreatif (Arumsari & Nursari, 2024). Melalui beberapa arah kebijakan industri dari pemerintah Indonesia dalam dua puluh tahun kedepan, subsektor industri kreatif ini mulai menjadi perhitungan dan menjadi kebijakan industri nasional. Peraturan Menteri Ekonomi Kreatif/Kepala Badan Ekonomi Kreatif Nomor 9 Tahun 2025 adalah regulasi yang mengatur tentang Penyelenggaraan Ekonomi Kreatif di Daerah di mana kebijakan ini sangat jelas untuk mendorong subsektor industri kreatif menjadi basis dalam pengembangan sektor industri nasional, salah satunya adalah fashion. Meskipun berbagai penelitian terdahulu telah menunjukkan efektivitas algoritma Apriori dalam konteks retail konvensional, penelitian yang secara spesifik mengkaji pola asosiasi produk fashion pada transaksi penjualan retail online masih perlu diperdalam. Retail online memiliki karakteristik yang berbeda dengan retail konvensional, termasuk volume transaksi yang lebih besar, variasi produk yang lebih beragam, serta perilaku konsumen yang cenderung lebih eksploratif dalam melakukan pembelian dari masa ke masa. Oleh karena itu, penelitian ini bertujuan untuk menemukan kumpulan produk

yang paling sering dibeli secara bersamaan oleh konsumen dalam transaksi retail online berdasarkan parameter *minimum support*, menghasilkan aturan hubungan antar produk yang kuat dengan mengukur nilai *confidence* dan *lift ratio* guna memahami keterkaitan antar item, memetakan kebiasaan belanja konsumen pada platform digital yang memiliki karakteristik volume data besar dan variasi produk yang beragam, serta merumuskan usulan taktis bagi pengelola retail online, seperti strategi penawaran paket (*bundling*), penempatan produk pada website, atau sistem rekomendasi produk khususnya produk fashion untuk meningkatkan profitabilitas.

METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini merupakan serangkaian tahapan yang terstruktur dan sistematis untuk mengidentifikasi pola asosiasi produk pada transaksi penjualan retail online menggunakan algoritma Apriori. Metode penelitian ini mencakup tahapan identifikasi masalah, pengumpulan dan prapemrosesan data, pemodelan data menggunakan algoritma Apriori, pembentukan aturan asosiasi, serta evaluasi dan interpretasi hasil. Berikut ini adalah penjelasan rinci mengenai setiap tahapan metode penelitian yang dilakukan.

Data penelitian diperoleh dari dataset *Fashion Retail Sales* yang berisi transaksi penjualan produk fashion secara online dengan jumlah 3.400 baris data. Populasi penelitian mencakup seluruh transaksi yang ada dalam dataset tersebut. Fokus penelitian dibatasi pada kategori produk fashion wanita agar hasil analisis lebih terarah. Dengan demikian, sampel penelitian yang dianalisis adalah transaksi konsumen yang memenuhi kriteria tersebut.

Tahap prapemrosesan data merupakan langkah penting yang dilakukan sebelum data dapat diproses menggunakan algoritma Apriori. Prapemrosesan data bertujuan untuk memastikan kualitas data yang baik sehingga hasil analisis yang diperoleh akurat dan dapat diandalkan (Saad et al., 2025). Tahap ini meliputi beberapa sub-proses sebagai berikut:

- a. Pembersihan data (Data Cleaning) dilakukan untuk mengatasi berbagai permasalahan pada data mentah, termasuk penanganan missing value, penghapusan data duplikat, dan koreksi inkonsistensi format. Missing value pada atribut-atribut penting diatasi menggunakan teknik imputation, yaitu pengisian nilai yang hilang dengan nilai modus atau nilai yang paling sering muncul pada atribut tersebut. Data duplikat yang terdeteksi harus dihapus untuk menghindari bias dalam hasil analisis. Selain itu, format data yang tidak konsisten perlu diseragamkan agar proses pemodelan dapat berjalan dengan lancar. Proses pembersihan data ini memastikan bahwa hanya data yang bersih dan konsisten yang akan diproses pada tahapan selanjutnya.
- b. Transformasi data dilakukan untuk mengubah data ke dalam format yang sesuai agar dapat diproses oleh algoritma Apriori. Proses transformasi meliputi pengelompokan data transaksi berdasarkan nomor transaksi dan tanggal, sehingga setiap baris data merepresentasikan satu nota transaksi yang berisi seluruh produk yang dibeli dalam transaksi tersebut. Setelah pengelompokan dilakukan, data ditransformasikan ke dalam format biner, di mana nilai 1 menunjukkan bahwa suatu produk dibeli dalam transaksi tersebut dan nilai 0 menunjukkan bahwa produk tersebut tidak dibeli. Format biner ini merupakan format standar yang diperlukan oleh algoritma Apriori untuk melakukan perhitungan nilai support dan confidence secara efisien.

Langkah-langkah dalam algoritma Apriori adalah sebagai berikut: (1) Pembentukan kandidat itemset, yaitu kandidat itemset dibentuk berdasarkan kombinasi dari itemset berukuran $(k-1)$ yang telah diperoleh pada iterasi sebelumnya. Salah satu karakteristik algoritma Apriori adalah dilakukannya pemangkasan terhadap kandidat k -itemset apabila ada subset-nya yang tidak termasuk dalam kumpulan item dengan frekuensi tinggi pada tahap sebelumnya; (2) Penghitungan support setiap kandidat k -itemset, yaitu untuk menentukan nilai support setiap kandidat, dilakukan pencarian terhadap seluruh basis data guna menghitung jumlah transaksi yang mengandung seluruh item dalam kandidat tersebut; (3) Penentuan pola frekuensi tinggi, yaitu pola frekuensi tinggi yang terdiri dari k item (k -itemset) ditentukan dari kandidat k -itemset yang memiliki nilai support melebihi ambang batas minimum support yang telah ditetapkan; (4) Penghentian proses atau iterasi berikutnya, yaitu jika tidak ditemukan pola frekuensi tinggi yang baru, maka proses dihentikan, namun jika masih ada pola yang memenuhi kriteria, nilai k ditingkatkan satu dan proses kembali dilanjutkan ke tahap pertama (Saad et al., 2025)(Azizah & Atmojo, 2025).

Eksplorasi awal dilakukan menggunakan analisis statistik deskriptif untuk mengetahui jumlah transaksi, jumlah produk unik, serta frekuensi kemunculan item. Hasil eksplorasi digunakan sebagai dasar dalam penentuan parameter algoritma Apriori. Tahapan algoritma Apriori secara garis besar terbagi menjadi dua proses utama: pencarian pola item yang sering muncul (Frequent Itemset Generation) dan pembentukan aturan hubungan antar item (Association Rule Generation) (Wijaya et al., 2024). Analisis item yang sering muncul (Frequent Itemset Generation) dilakukan untuk mencari kombinasi item yang memenuhi syarat minimal nilai support.

Penting atau tidaknya suatu aturan asosiasi dapat diketahui dengan beberapa parameter yaitu nilai support, confidence dan lift ratio (Khoiruzzidan & Iswari, 2024). Support adalah persentase dari jumlah kombinasi item dalam dataset. Nilai support untuk satu itemset dapat diperoleh dengan menggunakan rumus 1, nilai support dua item dapat diperoleh dengan menggunakan rumus 2, sedangkan nilai support tiga item dapat diperoleh dengan menggunakan rumus 3. Pencarian pola frekuensi tinggi akan dihentikan jika kombinasi tidak memenuhi nilai Support yang ditentukan.

$$\text{Support (A)} = \frac{\text{total transaksi mengandung A}}{\text{total transaksi}} \times 100\% \dots\dots\dots (1)$$

$$\text{Support (A} \cap \text{B)} = \frac{\text{total transaksi mengandung A dan B}}{\text{total transaksi}} \times 100\% \dots\dots\dots (2)$$

$$\text{Support (A, B, C)} = \frac{\text{total transaksi mengandung A,B dan C}}{\text{total transaksi}} \times 100\% \dots\dots\dots (3)$$

Setelah semua pola frekuensi tinggi ditemukan, selanjutnya dicari aturan asosiatif yang memenuhi syarat minimum kepercayaan yaitu 50% dengan menghitung aturan kepercayaan asosiatif "Jika A maka B" ($A \rightarrow B$). Nilai confidence dari aturan $A \rightarrow B$ dapat diperoleh menggunakan rumus berikut (Khoiruzzidan & Iswari, 2024):

$$\text{Confidence} = P(B | A) = \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi mengandung A}} \dots\dots\dots (4)$$

Pembentukan aturan asosiasi akhir diurutkan berdasarkan minimum support dan minimum confidence.

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining melalui analisis pola asosiasi untuk menemukan keterkaitan antar produk pada transaksi penjualan retail online. Tujuannya adalah untuk mengidentifikasi pola hubungan antar produk yang sering dibeli bersamaan dalam transaksi penjualan retail secara online. Data yang digunakan bersumber dari situs Kaggle dengan judul dataset "Fashion Retail Sales" yang berisi data transaksi penjualan produk fashion secara online. Dataset ini diunduh dalam format Excel (csv) dan memiliki total 3.400 baris data transaksi. Proses pembentukan model dilakukan secara sistematis dengan menetapkan parameter minimum support sebesar 7% dan minimum confidence sebesar 50%. Angka ini dipilih untuk memastikan bahwa aturan asosiasi yang dihasilkan memiliki validitas statistik yang kuat dan relevan secara bisnis. Dimulai dari pembentukan 1-itemset, penelitian ini menyaring produk tunggal yang memenuhi ambang batas frekuensi, yang kemudian dikembangkan menjadi 2-itemset untuk melihat pasangan produk yang kerap muncul bersamaan. Penggunaan operator W-Apriori di RapidMiner mempermudah proses komputasi ini, memastikan setiap kombinasi item dihitung secara akurat berdasarkan kemunculannya dalam ID transaksi yang unik.

Salah satu permasalahan yang dihadapi oleh pelaku retail online adalah bagaimana mengidentifikasi pola asosiasi antar produk yang sering dibeli secara bersamaan oleh konsumen. Pemahaman mengenai pola asosiasi produk ini sangat penting karena dapat memberikan wawasan tentang hubungan antar item dalam transaksi pembelian. Informasi ini dapat dimanfaatkan untuk berbagai keperluan strategis, seperti penentuan penempatan produk yang tepat, penyusunan penawaran bundling, pemberian rekomendasi produk kepada pelanggan, serta pengelolaan stok inventori yang lebih efisien. Permasalahan utama ini berfokus pada upaya untuk mengidentifikasi pola hubungan antar produk yang sering dibeli bersamaan dalam transaksi penjualan retail secara online. Menemukan pola tersebut menjadi sangat penting karena dapat memberikan nilai strategis bagi perusahaan, terutama dalam menyusun strategi pemasaran yang lebih tepat sasaran, mengoptimalkan tata letak produk di platform penjualan, dan menyusun penawaran paket (*bundling*) yang sesuai dengan pola perilaku konsumen. Oleh karena itu, penerapan metode asosiasi dalam penelitian ini diharapkan mampu menyajikan

informasi yang mendalam dan bermanfaat untuk mendukung proses pengambilan keputusan bisnis secara lebih efektif dan berbasis data.

Data yang digunakan bersumber dari situs Kaggle dengan judul dataset "Fashion Retail Sales" yang berisi data transaksi penjualan produk fashion secara online. Dataset ini diunduh dalam format Excel (csv) dan memiliki total 3.400 baris data transaksi. Setelah dilakukan proses *data cleaning* berupa penghapusan data duplikat, pengisian nilai kosong, dan transformasi ke bentuk pivot biner, jumlah data yang dapat digunakan adalah sebanyak 166 transaksi. Hal ini terjadi karena dalam dataset, satu pelanggan yang melakukan transaksi sering kali membeli lebih dari satu produk.

Pada data awal, setiap pembelian produk dicatat dalam baris yang terpisah, meskipun berasal dari transaksi yang sama. Melalui proses pivot, seluruh produk yang termasuk dalam satu ID transaksi dikonsolidasikan ke dalam satu baris data. Artinya, setiap ID transaksi yang unik akan menjadi satu baris dalam tabel pivot, sedangkan kolom-kolomnya berisi produk-produk yang dibeli dalam transaksi tersebut. Hasil dari proses pivot akan menghasilkan tabel dengan format sebagai berikut: angka 1 menunjukkan bahwa produk tersebut dibeli dalam transaksi yang bersangkutan, angka 0 menunjukkan bahwa produk tersebut tidak dibeli dalam transaksi yang bersangkutan.

Setelah proses pivot selesai, langkah selanjutnya adalah memberikan nama atribut yang lebih rapi dan seragam menggunakan operator Rename by Replacing di Rapid Miner. Pada hasil pivot, biasanya nama atribut yang terbentuk otomatis akan panjang dan tidak terformat dengan baik, contohnya: count(Customer Reference ID)_Sweater. Agar nama atribut lebih sederhana dan mudah dibaca, proses Rename by Replacing dilakukan untuk menghapus bagian count (Customer Reference ID)_ sehingga hanya tersisa nama produk seperti Sweater, Raincoat, Backpack, dan sebagainya.

Selanjutnya adalah menangani missing value pada data, yang ditandai dengan tanda "?". Missing value ini diganti menjadi angka 0 menggunakan operator Replace Missing Values agar data dapat diproses pada tahap selanjutnya. Data hasil pivot yang masih bertipe numerik kemudian diubah menjadi tipe binomial menggunakan operator Numerical to Binomial. Hal ini penting agar data dapat diolah menggunakan algoritma Apriori yang memerlukan input bertipe binomial. Langkah terakhir pada tahap preprocessing adalah memberikan peran pada atribut Customer Reference ID sebagai ID dengan menggunakan operator Set Role. Hal ini bertujuan agar Rapid Miner mengenali Customer Reference ID sebagai identitas unik setiap transaksi.

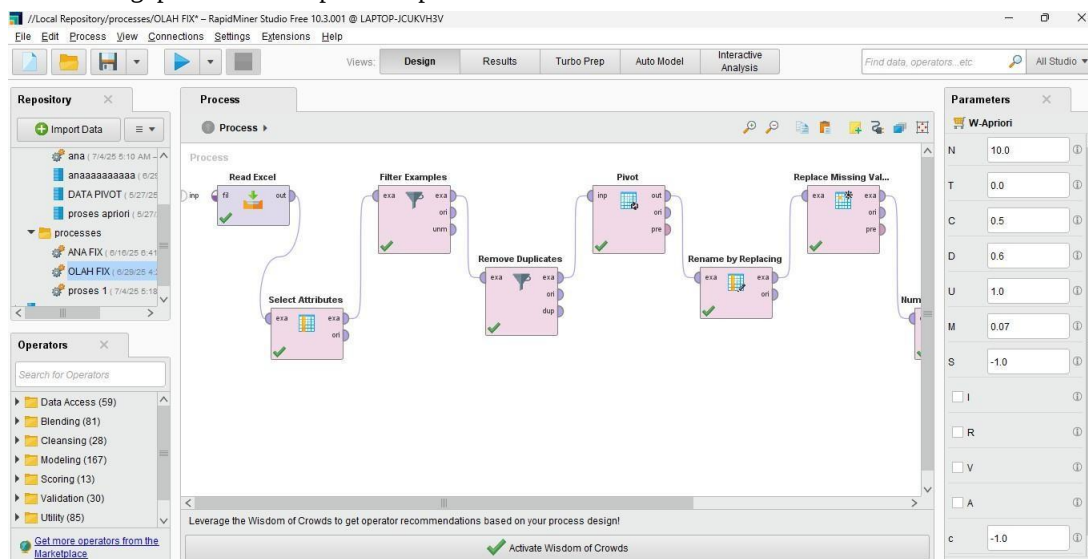
Setelah data dibersihkan, dilakukan eksplorasi data untuk memahami karakteristik umum dari dataset. Eksplorasi ini bertujuan untuk memberikan gambaran awal sebelum proses pembentukan model. Hasil eksplorasi menunjukkan informasi sebagai berikut: Jumlah total transaksi: 166, Jumlah produk unik: 50 item, produk dengan tingkat frekuensi kemunculan tertinggi adalah: Belt sebanyak 68 kali, Skirt sebanyak 68 kali, Tie sebanyak 68 kali, Pajamas sebanyak 67 kali, Pants sebanyak 66 kali

Informasi ini mengindikasikan bahwa produk-produk tersebut cenderung dibeli oleh konsumen, baik secara individu maupun bersama dengan produk lainnya. Produk-produk dengan frekuensi tinggi ini menjadi kandidat kuat dalam pembentukan pola asosiasi. Selain itu, distribusi frekuensi kemunculan item dapat digunakan untuk menentukan nilai minimum support yang ideal agar proses pembentukan model menghasilkan frequent itemset yang relevan, namun tidak terlalu banyak atau terlalu sedikit. Data hasil eksplorasi ini kemudian menjadi dasar dalam proses selanjutnya, yaitu pembentukan model asosiasi menggunakan algoritma Apriori yang diawali dengan perhitungan kombinasi satu item (C1), dua item (C2), hingga tiga item (C3).

Seleksi Fitur (Feature Selection). Feature selection diterapkan pada fase ini untuk menyaring atribut yang signifikan, yaitu pemilihan atribut yang dianggap relevan untuk mendukung analisis pola asosiasi antar produk. Dari keseluruhan atribut yang terdapat dalam dataset, hanya dua atribut yang dipertahankan, yaitu: Customer Reference ID, digunakan untuk mengidentifikasi transaksi secara individual berdasarkan pelanggan. Item Purchased, digunakan untuk mengetahui produk yang dibeli dalam setiap transaksi. Sementara itu, atribut lain seperti Purchase Amount (USD), Review Rating, dan Payment Method tidak disertakan karena tidak memiliki **keterkaitan** langsung dengan tujuan analisis asosiasi produk. Dengan melakukan seleksi fitur ini, proses analisis menjadi lebih terfokus, sederhana, dan efisien, tanpa mengurangi kualitas informasi yang dibutuhkan dalam pembentukan pola asosiasi.

Pembentukan Model (Penerapan Algoritma Apriori). Setelah melalui tahapan preprocessing, eksplorasi data, dan seleksi fitur, langkah selanjutnya adalah melakukan pembentukan model guna menemukan pola asosiasi antar produk pada transaksi penjualan retail online. Model ini dibangun menggunakan algoritma Apriori dengan bantuan aplikasi Rapid Miner. Proses pembentukan model sepenuhnya dilakukan di Rapid Miner. Dataset yang telah diproses sebelumnya kemudian diubah menjadi pivot table dalam format biner menggunakan operator Pivot. Setelah data berhasil dipivot, atribut Customer Reference ID disetel sebagai identitas transaksi unik menggunakan operator Set Role, agar setiap transaksi dapat dikenali secara individual oleh sistem. Selanjutnya, seluruh atribut produk yang masih bertipe numerik diubah menjadi tipe binomial menggunakan operator Numerical to Binomial, agar sesuai dengan format data yang dibutuhkan oleh algoritma Apriori. Proses pembentukan model

dilakukan dengan operator W-Apriori. Parameter yang diterapkan dalam proses pembentukan model ini antara lain sebagai berikut: Minimum Support: 0.07 (7%), berarti aturan yang terbentuk harus muncul pada minimal 7% dari total transaksi. Minimum Confidence: 0.5 (50%), berarti aturan yang dihasilkan memiliki tingkat keyakinan minimal 50% bahwa jika suatu produk dibeli, maka produk lainnya kemungkinan besar juga akan dibeli dalam transaksi yang sama. Penetapan parameter tersebut dimaksudkan untuk mengidentifikasi kombinasi produk yang kerap muncul serta memiliki nilai confidence yang tinggi. Tahap ini dilakukan untuk menghasilkan himpunan item frekuensi tinggi (frequent itemset) serta membangun aturan asosiasi yang dapat digunakan sebagai dasar dalam merumuskan strategi pemasaran dan penataan produk.



Sumber: Mufida, 2026

Gambar 1 Pembentukan Model Menggunakan Operator W-Apriori di RapidMiner

Proses pembentukan 1-itemset (C1) dilakukan untuk mengetahui frekuensi kemunculan masing-masing produk secara individual dalam transaksi. Pada tahap ini, ditetapkan nilai *minimum support* sebesar 0,07 (7%). Produk-produk yang tidak mencapai ambang batas tersebut akan dikeluarkan dari proses analisis selanjutnya. Nilai *support* untuk masing-masing produk dihitung menggunakan Rumus (1), yaitu dengan membandingkan jumlah transaksi yang mengandung produk tertentu terhadap total keseluruhan transaksi. Berikut adalah daftar 1-itemset yang memenuhi kriteria *minimum support* dan digunakan dalam pembentukan itemset selanjutnya.

Pembentukan 2-itemset (C2) dilakukan untuk mengidentifikasi pasangan produk yang sering muncul secara bersamaan dalam satu transaksi. Pada tahap ini, digunakan ambang batas *minimum support* sebesar 0,07 (7%). Nilai *support* dari setiap kombinasi dua item dihitung menggunakan Rumus (2), yaitu dengan membandingkan jumlah transaksi yang memuat kombinasi tersebut terhadap total keseluruhan transaksi. Kombinasi item yang tidak memenuhi nilai *minimum support* dieliminasi dan tidak dilanjutkan ke proses pembentukan itemset pada tahap berikutnya. Tahap selanjutnya dalam analisis melibatkan pembentukan 3-itemset untuk mengidentifikasi kelompok yang terdiri dari tiga produk yang sering dibeli bersamaan dalam satu transaksi. Proses ini dilakukan dengan menggabungkan produk-produk yang sebelumnya telah memenuhi nilai *minimum support* pada tahap 2-itemset. Tujuan dari pembentukan 3-itemset adalah untuk mengidentifikasi pola pembelian yang lebih kompleks dan mendalam, dengan melibatkan tiga item dalam satu transaksi. Nilai *support* dihitung menggunakan Rumus (2). Setelah pola-pola dengan frekuensi tinggi berhasil diidentifikasi, langkah berikutnya adalah membentuk aturan asosiasi yang memenuhi ambang *minimum confidence* sebesar 50%. Nilai *confidence* untuk asosiasi $A \rightarrow B$ dihitung menggunakan Rumus (3). Aturan yang memiliki nilai *confidence* di bawah ambang *minimum* akan disingkirkan dari analisis. Rincian perhitungan nilai *confidence* antar produk dapat dilihat pada Tabel IV.4.

Tabel 4 Aturan Asosiasi

No	Aturan	Confidence
1	Jika membeli Bowtie dan Ponco maka akan membeli Belt	12/15 80%
2	Jika membeli Sandals dan Skirt maka akan membeli Belt	12/15 80%
3	Jika membeli Sneakers dan Tunic maka akan membeli Skirt	12/15 80%
4	Jika membeli Belt dan Dress maka akan membeli Skirt	14/18 78%

5	Jika membeli Jacket dan Sneakers maka akan membeli Belt	13/17	76%
6	Jika membeli Kimono dan Shorts maka akan membeli Tanktop	16/21	76%
7	Jika membeli Backpack dan Sweater maka akan membeli Raincoat	12/16	75%
8	Jika membeli Leggings dan Sweater maka akan membeli Camisole	12/16	75%
9	Jika membeli Socks dan Sweater maka akan membeli Camisole	12/16	75%
10	Jika membeli Jacket dan Raincoat maka akan membeli Loafers	12/16	75%

Sumber: Mufida, 2026

Pengujian terhadap model dilakukan dengan menggunakan tools aplikasi Rapid Miner. Parameter yang digunakan dalam pengujian ini adalah minimum support sebesar 0,07 (7%) dan minimum confidence sebesar 0,5 (50%). Dataset yang dianalisis terdiri dari 166 transaksi yang sebelumnya telah melalui proses preprocessing dan diubah ke dalam bentuk format biner. Tujuan dari pengujian ini adalah untuk menemukan pola asosiasi dalam pembelian produk secara bersamaan pada setiap transaksi.

Dalam proses pengujian, ditetapkan nilai ambang minimum support sebesar 0,07 (7%) dan minimum confidence sebesar 0,5 (50%) yang telah disesuaikan dengan karakteristik dataset yang terdiri dari 166 transaksi. Parameter tersebut dipilih untuk menangkap pola pembelian produk yang cukup sering muncul, tanpa menghilangkan potensi kombinasi yang bernilai informatif. Dari hasil pemrosesan algoritma apriori menghasilkan: 50 itemset tunggal (L1), 1174 kombinasi dua item (L2), dan 441 kombinasi tiga item (L3). Dari berbagai kombinasi yang dihasilkan, teridentifikasi 10 aturan asosiasi unggulan dengan tiga nilai confidence tertinggi, yaitu:

- 1 Jika konsumen membeli Bowtie dan Poncho, maka kemungkinan besar juga membeli Belt, dengan tingkat support sebesar 7,2% dan confidence sebesar 80%.
- 2 Jika konsumen membeli Sandals dan Skirt, maka kemungkinan besar juga akan membeli Belt, dengan tingkat support sebesar 7.2% dan confidence sebesar 80%.
- 3 Jika membeli Sneakers dan Tunic, maka kemungkinan besar juga akan membeli Skirt, dengan tingkat support sebesar 7.2% dan confidence sebesar 80%.

Tahap selanjutnya adalah melakukan evaluasi model untuk menilai efektivitas algoritma Apriori dalam menemukan pola hubungan antara produk dalam data transaksi. Pada tahap ini, model dikembangkan dengan menetapkan nilai minimum support sebesar 0,07 (7%) dan minimum confidence sebesar 0,5 (50%). Hasil pengolahan data menggunakan Rapid Miner menghasilkan 10 aturan asosiasi yang memenuhi kriteria yang telah ditetapkan. Berikut adalah hasil aturan asosiasi yang terbentuk berdasarkan output dari RapidMiner: Konsumen yang membeli Bowtie dan Poncho cenderung juga membeli Belt, dengan nilai support 7.2% dan confidence 80%. Artinya, sebanyak 7.2% dari seluruh transaksi mencakup ketiga produk tersebut, dan dari seluruh transaksi yang mengandung Bowtie dan Poncho, 80% di antaranya juga menyertakan Belt. Dari hasil tersebut, dapat disimpulkan bahwa model Apriori mampu mengidentifikasi pola keterkaitan produk yang cukup kuat dan relevan. Aturan-aturan ini berpotensi digunakan sebagai dasar dalam menyusun strategi bisnis, misalnya dalam penyusunan rekomendasi produk, pengelompokan barang di toko, hingga promosi bundling yang lebih tepat sasaran.

Dengan ditemukannya pola hubungan antar produk, perusahaan dapat menggunakannya untuk menyusun strategi pemasaran yang lebih efektif, seperti penempatan posisi produk yang saling berkaitan secara berdekatan (*product placement*) atau pembuatan paket bundling yang menarik. Implementasi algoritma Apriori ini membuktikan bahwa pengolahan data transaksi yang tepat dapat mengubah sekumpulan catatan penjualan menjadi informasi strategis yang mampu meningkatkan efisiensi operasional dan kepuasan pelanggan di platform retail online. Data ini dapat digunakan sebagai dasar untuk penarikan kesimpulan serta rekomendasi strategis dalam hal pemasaran atau pengelompokan produk.

KESIMPULAN

Algoritma Apriori terbukti efektif dalam memetakan pola perilaku belanja konsumen pada dataset *Fashion Retail Sales* dengan parameter *minimum support* 7% dan *minimum confidence* 50%. Produk yang paling sering dibeli (*frequent items*) oleh pelanggan adalah Belt, Skirt, Tie, Pajamas, dan Pants. Ditemukan 10 aturan asosiasi utama, dengan 3 aturan teratas yang memiliki nilai *confidence* sebesar 80%, jika membeli Bowtie dan Poncho, maka akan membeli Belt, jika membeli Sandals dan Skirt, maka akan membeli Belt, dan jika membeli Sneakers dan Tunic, maka akan membeli Skirt. Proses prapemrosesan (seperti *pivot* dan *cleaning*) berhasil mengubah data mentah sebanyak 3.400 baris menjadi 166 transaksi unik yang siap dianalisis untuk kebutuhan strategis.

Berdasarkan temuan pola asosiasi di atas, strategi bisnis yang direkomendasikan adalah membuat paket promosi atau diskon khusus untuk produk yang memiliki hubungan kuat, misalnya paket "Aksesori Lengkap" (Bowtie + Poncho + Belt) atau paket "Gaya Kasual" (Sneakers + Tunic + Skirt). Pada platform retail online,

produk-produk yang sering dibeli bersamaan dapat ditampilkan dalam satu halaman atau di bagian "Sering dibeli bersamaan" (*Frequently Bought Together*) untuk memudahkan konsumen. Penulis merekomendasikan untuk mengintegrasikan hasil aturan asosiasi ke dalam sistem rekomendasi otomatis pada website. Jika pelanggan memasukkan *Sandals* dan *Skirt* ke keranjang, sistem dapat secara otomatis menyarankan *Belt*. Memastikan ketersediaan stok produk pendukung (seperti *Belt* dan *Skirt*) selalu terjaga saat produk pemicunya (seperti *Bowtie* atau *Tunic*) sedang dalam masa promosi atau tren tinggi. Mengingat perilaku konsumen di platform digital cenderung dinamis dan eksploratif, penelitian ini sebaiknya dilakukan secara rutin untuk memperbarui pola asosiasi seiring perubahan tren fashion.

Hasil penelitian ini dapat diintegrasikan langsung ke dalam ekosistem retail melalui otomatisasi sistem rekomendasi, *bundling* promosi dinamis, dan optimalisasi manajemen inventaris, serta berpotensi dikembangkan lebih lanjut menggunakan *Sequential Pattern Mining*, algoritma alternatif, dan metode *clustering*. Penelitian lanjutan dapat difokuskan pada analisis tren *fashion* musiman dan pemetaan perilaku belanja berbasis segmentasi demografi untuk meningkatkan efisiensi komputasi dan akurasi prediksi stok.

REFERENSI

- Ananda, Y., Sembiring, B., & Sembiring, E. A. (2023). *Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang*. 1(1), 1–8.
- Arumsari, A., & Nursari, F. (2024). *PELUANG PENGEMBANGAN PRODUK RAMAH LINGKUNGAN PADA*. 7(April), 11–16.
- Azizah, Q., & Atmojo, W. T. (2025). *Penerapan Algoritma Apriori Dalam Analisis Pola Keanggotaan Fitness Untuk Menentukan Strategi Promosi*. 19(1), 30–41.
- Brighton, K., & Hariyanto, S. (2024). *Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik*. 7(1). <https://doi.org/10.32877/bt.v7i1.1417>
- Fuadi, A. N., Bhakti, H., & Premana, A. (2024). *RITEL DMART MENGGUNAKAN ALGORITMA APRIORI*. 12(3), 3269–3277.
- Hanapi, A., & Sari, R. (2023). *Penerapan Algoritma Apriori Untuk Rekomendasi Produk Bagi Pelanggan Toko Online Berbasis Website*. 5(1), 51–60.
- Jufri, F. R., Defit, S., & Nurcahyo, G. W. (2024). *Jurnal KomtekInfo Penerapan Algoritma Apriori dalam Menentukan Pola Penjualan*. 11(4), 363–370. <https://doi.org/10.35134/komtekinfo.v11i4.583>
- Khoiruzzidan, I., & Iswari, L. (2024). *Jurnal Sains , Nalar , dan Aplikasi Teknologi Informasi*. 3(2), 50–57. <https://doi.org/10.20885/snati.v3.i2.31>
- Mubarok, A., Suherman, H. D., Ramdhani, Y., & Topiq, S. (2019). *Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Dengan Metode TOPSIS*. *Jurnal Informatika*, 6(1), 37–46. <https://doi.org/10.31311/ji.v6i1.4739>
- Nur, A. M., Bahtiar, H., Suhartini, Fathurrahman, I., & Muzarrofah, L. R. (2025). *Penerapan Algoritma Apriori Dalam Mengidentifikasi Pola Perilaku Belanja Konsumen*. 8(2).
- Nurdiansyah, R. M., Rahmawati, A. S., Prayitno, S., & Rifani, D. A. (2024). *Analisis Pemilihan Platform E-Commerce Menggunakan Metode AHP : Studi Kasus Perilaku Konsumen Online*. 3(2), 85–101.
- Saad, F. S., Azizah, N., & Yani, Z. (2025). *Penerapan Algoritma Apriori untuk Menentukan Pola Permintaan Bahan Makanan Berdasarkan Data Transaksi*. 5–8.
- Safitry, D. L., Rosianti, N., Divyaning, E., Zidan, H., Arnecia, Z. J., Paryudi, I., Veritawati, I., Rezeki, S., & Nursari, C. (2025). *PEMASARAN PRODUK DI TOKO RETAIL X*. 9(1), 505–511.
- Sutoyo, M. N., Mardianto, & Paliling, A. (2024). *Implementasi Algoritma Apriori untuk Mengidentifikasi Tren Pemilihan Jurusan*. 10(2), 129–137.
- Tam, J., Issn, E., Issn, P., Hasna, A., Shafa, F., Perdana, A., Informasi, P. S., Dharma, S., & Metro, W. (2024). *PENERAPAN ALGORITMA APRIORI PADA DATA TRANSAKSI PENJUALAN DI TOKO BERKAH BERLIMPAH*. 15(1), 88–94.
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). *Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap)*. *JUSTIFY: Jurnal Sistem Informasi Ibrahimi*, 12(1), 51–61. <https://doi.org/10.35316/justify.v3i1.5335>
- Wijaya, N. S., Jajuli, M., & Dermawan, B. A. (2024). *Penerapan Algoritma K-Means Clustering Dalam Menentukan Daerah Prioritas Penanganan Kemiskinan Di Wilayah Jawa Timur*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8 (4), 7579–7584. <https://doi.org/10.36040/jati.v8i4.10248>
- Ananda, Y., Sembiring, B., & Sembiring, E. A. (2023). *Implementasi Data Mining Menggunakan Algoritma*

- Apriori Dalam Menentukan Persediaan Barang*. 1(1), 1–8.
- Arumsari, A., & Nursari, F. (2024). *PELUANG PENGEMBANGAN PRODUK RAMAH LINGKUNGAN PADA*. 7(April), 11–16.
- Azizah, Q., & Atmojo, W. T. (2025). *Penerapan Algoritma Apriori Dalam Analisis Pola Keanggotaan Fitness Untuk Menentukan Strategi Promosi*. 19(1), 30–41.
- Brighton, K., & Hariyanto, S. (2024). *Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik*. 7(1). <https://doi.org/10.32877/bt.v7i1.1417>
- Fuadi, A. N., Bhakti, H., & Premana, A. (2024). *RITEL DMART MENGGUNAKAN ALGORITMA APRIORI*. 12(3), 3269–3277.
- Hanapi, A., & Sari, R. (2023). *Penerapan Algoritma Apriori Untuk Rekomendasi Produk Bagi Pelanggan Toko Online Berbasis Website*. 5(1), 51–60.
- Jufri, F. R., Defit, S., & Nurcahyo, G. W. (2024). *Jurnal KomtekInfo Penerapan Algoritma Apriori dalam Menentukan Pola Penjualan*. 11(4), 363–370. <https://doi.org/10.35134/komtekinfo.v11i4.583>
- Khoiruzzidan, I., & Iswari, L. (2024). *Jurnal Sains , Nalar , dan Aplikasi Teknologi Informasi*. 3(2), 50–57. <https://doi.org/10.20885/snati.v3.i2.31>
- Mubarak, A., Suherman, H. D., Ramdhani, Y., & Topiq, S. (2019). *Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Dengan Metode TOPSIS*. *Jurnal Informatika*, 6(1), 37–46. <https://doi.org/10.31311/ji.v6i1.4739>
- Nur, A. M., Bahtiar, H., Suhartini, Fathurrahman, I., & Muzarrofah, L. R. (2025). *Penerapan Algoritma Apriori Dalam Mengidentifikasi Pola Perilaku Belanja Konsumen*. 8(2).
- Nurdiansyah, R. M., Rahmawati, A. S., Prayitno, S., & Rifani, D. A. (2024). *Analisis Pemilihan Platform E-Commerce Menggunakan Metode AHP : Studi Kasus Perilaku Konsumen Online*. 3(2), 85–101.
- Saad, F. S., Azizah, N., & Yani, Z. (2025). *Penerapan Algoritma Apriori untuk Menentukan Pola Permintaan Bahan Makanan Berdasarkan Data Transaksi*. 5–8.
- Safitry, D. L., Rosianti, N., Divyaning, E., Zidan, H., Arnecia, Z. J., Paryudi, I., Veritawati, I., Rezeki, S., & Nursari, C. (2025). *PEMASARAN PRODUK DI TOKO RETAIL X*. 9(1), 505–511.
- Sutoyo, M. N., Mardianto, & Paliling, A. (2024). *Implementasi Algoritma Apriori untuk Mengidentifikasi Tren Pemilihan Jurusan*. 10(2), 129–137.
- Tam, J., Issn, E., Issn, P., Hasna, A., Shafa, F., Perdana, A., Informasi, P. S., Dharma, S., & Metro, W. (2024). *PENERAPAN ALGORITMA APRIORI PADA DATA TRANSAKSI PENJUALAN DI TOKO BERKAH BERLIMPAH*. 15(1), 88–94.
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). *Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap)*. *JUSTIFY: Jurnal Sistem Informasi Ibrahmy*, 12(1), 51–61. <https://doi.org/10.35316/justify.v3i1.5335>
- Wijaya, N. S., Jajuli, M., & Dermawan, B. A. (2024). *Penerapan Algoritma K-Means Clustering Dalam Menentukan Daerah Prioritas Penanganan Kemiskinan Di Wilayah Jawa Timur*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8 (4), 7579–7584. <https://doi.org/10.36040/jati.v8i4.10248>