Implementasi Data Mining dengan Algoritma K-Medoids dan FP-Growth untuk Menentukan Asosiasi Itemset Produk

|  |  |
| --- | --- |
| **Diterima:**10 Juni 2024**Revisi:**10 Juli 2024**Terbit:**1 Agustus 2024 | **1\*Moch. Ghufron Ramadhani, 2Patmi Kasih, 3Intan Nur Farida***1-3Universitas Nusantara PGRI Kediri*1*ghufronramadhani6@gmsil.com*,*2fatkasih@gmail.com,3intanf@unpkediri.ac.id* |

**Abstrak**— Ceria Store, sebagai salah satu bisnis retail, menghadapi tantangan meningkatkan efisiensi operasional dalam menghadapi persaingan industri ritel yang ketat. Oleh karena itu, diperlukan pendekatan yang lebih mendalam dan berbasis data untuk mengatasi masalah ini. Penelitian ini menggunakan algoritma *K-Medoids* untuk clustering dan *FP-Growth* untuk analisis asosiasi itemset, menganalisis pola pembelian konsumen berdasarkan data penjualan dari November dan Desember 2023. Proses *clustering* menunjukkan adanya tiga kelompok penjualan berbeda (rendah, sedang, dan tinggi), untuk memperkecil data yang akan diasosiasikan. Beberapa rekomendasi produk ditemukan dari analisis asosiasi berupa 4 item Produk Yaitu Daster Kencana Ungu, Short Ivanova, Sport Bra, Andin Underwear, Atasan Lyra, Leging Adi Sp.

**Kata Kunci**—K-Medoids;FP-Growth;Asosiasi Itemset;Clustering;Pola Pembelian Produk

***Abstract***— *Ceria Store, as a retail business, faces the challenge of increasing operational efficiency in the face of tight retail industry competition. Therefore, a more in-depth and data-driven approach is needed to overcome this problem. This research uses the K-Medoids algorithm for clustering and FP-Growth for itemset association analysis, analyzing consumer purchasing patterns based on sales data from November and December 2023. The clustering process shows the existence of three different sales groups (low, medium and high), to minimize data to be associated. Several product recommendations were found from the association analysis in the form of 4 product items, namely Daster Kencana Ungu, Short Ivanova, Sport Bra, Andin Underwear, Atasan Lyra, Leging Adi Sp.*

***Keywords****—K-Medoids;FP-Growth;Itemset Association;Clustering;Product Purchasing Patterns*

This is an open access article under the CC BY-SA License.

# A picture containing text, clipart  Description automatically generated

***Penulis Korespondensi:***

Nama Penulis, Moch. Ghufron Ramadhani

Departemen Penulis, Fakultas Teknik dan Ilmu Komputer

Institusi Penulis,Universitas Nusantara Kediri

Email: Email Penulis ghufronramadhani6@gmail.com
ID Orcid: [<https://orcid.org/register>]
Handphone: 085704570641



# **I. PENDAHULUAN**

 Dalam era digital yang semakin maju, industri retail menghadapi tantangan yang semakin kompleks dan kompetitif. Perubahan perilaku konsumen yang dipengaruhi oleh pesatnya perkembangan teknologi dan transformasi digital, khususnya dalam penjualan *online* [1], menuntut para pelaku bisnis retail untuk terus berinovasi dan memaksimalkan potensi yang ada. Ceria Store, sebagai salah satu bisnis retail pakaian di Kecamatan Gondang Kabupaten Nganjuk, Jawa Timur, berupaya untuk tetap relevan dan berdaya saing di tengah dinamika tersebut.

 Ceria Store menghadapi berbagai tantangan yang umum dihadapi oleh bisnis retail, seperti kesulitan dalam menentukan strategi pemasaran yang efektif dan manajemen stok yang optimal. Ketidaksesuaian produk yang ditawarkan dengan preferensi konsumen serta strategi pemasaran yang kurang efektif dapat menyebabkan penurunan penjualan dan peningkatan tingkat retur barang. Oleh karena itu, diperlukan pendekatan yang lebih mendalam dan berbasis data untuk mengatasi masalah ini.

 Penelitian ini bertujuan untuk menggunakan algoritma *K-Medoids* untuk clustering dan *FP-Growth* untuk *asosiasi itemset*. *K-Medoids* menggunakan objek yang paling terpusat dalam sebuah *cluster* [2], sedangkan *FP-Growth* menggunakan struktur data *FP-Tree* bersamaan dengan algoritma *FP-Growth* untuk menemukan *frequent itemset* [3], yang merupakan himpunan data yang paling sering muncul dalam suatu dataset. Dengan pendekatan ini, diharapkan dapat menemukan itemset produk yang paling diminati konsumen, menentukan strategi pemasaran yang lebih efektif berdasarkan pola penjualan yang telah dianalisis, menemukan pola asosiasi antar produk, dan mengidentifikasi kelompok barang dengan perilaku penjualan serupa.

Penelitian ini didasarkan pada beberapa penelitian sebelumnya yang telah menunjukkan efektivitas kombinasi algoritma *K-Medoids* dan *FP-Growth*. Wiwit Agus Triyanto dan Vincent Suhartono [4] menemukan bahwa kombinasi ini menghasilkan aturan *asosiasi* yang lebih akurat dibandingkan algoritma *Apriori*. Imaduddin Syukra dan Hidayat Asad [5] menunjukkan bahwa *K-Medoids* dan *FP-Growth* dapat digunakan untuk mengelompokkan produk dan menghasilkan rekomendasi yang sesuai dengan kebutuhan konsumen. Muhammad Imam Ghozali dan Ehwan Ri’fan Zaenal [6] menemukan bahwa kombinasi *FP-Growth*, *Self-Organizing Map (SOM*), dan *K-Medoids* dapat menganalisis pola belanja dan menghasilkan rekomendasi produk yang lebih efektif. Lalu penelitian [7] algoritma *apriori* digunakan sebagai proses utama dari analisa pola keterjangkitan penyakit. algoritma *apriori* untuk analisa pola dengan menggunakan minimum *support* untuk menemukan aturan *asosiasi.* Septiandy A.P.P dkk [8] menggunakan algoritma *apriori* yang menggunakan analisis 2 nilai penting yaitu *support* dan *confidence*, untuk menemukan frequent itemset yang dijalankan pada sekumpulan data

# **II. METODE**

1. Prosedur Penelitian

Adapun langkah-langkah untuk metode penelitian sebagai berikut:

* 1. Analisis Kebutuhan

Identifikasi kebutuhan yang diperlukan untuk penelitian dan sistem yang akan dibuat.

* 1. Desain Sistem

Perancangan antarmuka sistem untuk mendukung pengembangan sistem pendukung pola penjualan.

* 1. Implementasi

Pembuatan itemset pola penjualan setelah merancang sistem, menggunakan *PHP, CSS*, dan bahasa pemrograman lainnya. *Database MySQL* diimplementasikan.

* 1. Pengujian Program

Memastikan output sesuai dengan kebutuhan dan metode yang direncanakan dengan black box testing, menguji logika dan fungsionalitas sistem.

* 1. Pemeliharaan

Memastikan perangkat lunak tetap sesuai harapan yang ditetapkan dengan memperbaiki kode dan melakukan pengujian ulang jika diperlukan.

1. *K-Medoid*

Algoritma *K-Medoid* untuk perhitungan dan mengelompokan data, tahapan-tahapan yang dilakukan[9] yaitu:

Menentukan jumlah cluster:

Menghitung jarak data dengan pusat medoid;

Menentukan jarak terdekat dari setiap data;

Menentukan nilai medoid/titik pusat baru;

Menghitung jarak data dengan medoid baru;

Menentukan jarak terdekat dari setiap data;

Menghitung jumlah kedekatan;

Perbandingan jumlah kedekatan

1. *FP-Growth*

Algoritma *FP-Growth* yang digunakan untuk menemukan pola penjualan berupa itemset.

Lalu, tahap-tahap dalam proses algoritma *FP-Growth*[10] sebagai berikut:

1. Tahap FP-Tree

Pada tahap ini, jumlah *support* untuk setiap item dalam basis pola bersyarat dijumlahkan. Setiap item dengan dukungan yang memenuhi atau melebihi minimum dua akan dihasilkan menggunakan *conditional FP-Tree*. Rumus untuk *support* adalah:

$$support=P ∩ B=\frac{jumlah transaksi yang memuat A dan B}{Jumlah transaksi}$$

1. Tahap Pattern Base

*Pattern Base* adalah *sub-database* yang terdiri dari *prefix path* dan *suffix pattern* untuk menemukan *frequent* itemset. Pembentukan aturan dilakukan dari subset pada *frequent pattern* yang memenuhi nilai *support*. Nilai *confidence* dari aturan yang terbentuk dihitung menggunakan rumus:

$$Confidence=P\left(\frac{B}{A}\right)=\frac{Support (P ∩ B)}{P(A)}$$

1. Tahap Pencarian (*Frequent Itemset*)

*FP-Tree* menghasilkan *frequent itemset* dengan menggabungkan item pada setiap lintasan tunggal. Jika tidak*, FP-Growth* dipanggil secara *rekursif*. Terakhir, dilakukan perhitungan *lift ratio* untuk setiap aturan yang terbentuk menggunakan rumus:

$$Lift Ratio=\frac{Support (P ∩ B)}{Support\left(A\right)\*Support(B)}$$

1. Perancangan Sistem
2. *Squence Diagram*

Gambar 2 *Sequence Diagram* menunjukkan alur eksekusi dari proses *clustering* dan *asosiasi* dalam sistem saling berinteraksi.



Gambar 1. Squence Diagram

1. Kebutuhan Data

Pada gambar 3 data yang digunakan berupa data mentah penjualan di Ceria Store selama Juni dan Mei 2024, dengan 3085 transaksi dan 342 produk. Data akan diproses ke dalam metode *clustering K-medoid* dan selanjutnya diproses untuk menemukan nilai asosiasi menggunakan *FP-Growth.*



Gambar 2. Data Penjualan

1. Proses *Clustering*

proses clustering dilakukan dengan membagi data menjadi tiga kluster. kluster 0 dengan tingkat penjualan rendah, kluster 1 dengan tingkat penjualan sedang, dan kluster 2 dengan tingkat penjualan tinggi. Langkah ini bertujuan untuk memperkecil data yang akan diasosiasikan sehingga dapat menghasilkan pola yang lebih akurat.

1. Proses Asosiasi

proses *asosiasi* menggunakan metode *FP-Growth* dengan menetapkan minimum *support* dan *confidence* sebagai batas ambang atau *threshold* untuk menentukan aturan asosiasi terbaik.

# **III. HASIL DAN PEMBAHASAN**

1. Hasil Rekomendasi Produk

Setelah terbentuk *cluster* akan dilakukan proses asosiasi dengan algoritma *FP-Growth* mendapatkan rekomendasi produk untuk pelanggan, dengan item *Premises* Daster Kecana Ungu, dengan *Conclusion* Short Ivanova memiliki nilai *Support* sebesar0.18, *Confidence* Sebesar 0.8, dan *Lift Ratio* sebesar 3.68. lalu item *Premises* Short Ivanova dengan *Conclusion* Daster Kencana Ungu memiliki nilai *Support* sebesar0.18, *Confidence* Sebesar 0.85, dan *Lift Ratio* sebesar 3.68. item *Premises* Sport Bra dengan *Conclusion* Andin Underwear memiliki nilai *Support* sebesar0.05, *Confidence* Sebesar 5, dan *Lift Ratio* sebesar 4.12. item *Premises* Andin Underwear, dengan *Conclusion* Sport Bra memiliki nilai *Support* sebesar0.5, *Confidence* Sebesar 0.45, dan *Lift Ratio* sebesar 4.12. itemset *Premises* Atasan Lyra dengan Conclusion Leging Adi Sp memiliki nilai *Support* sebesar 0.11, dan *Confidence* 0.87, dan *Lift Ratio* sebesar 5.2. itemset *Premises* Leging Adi Sp dengan Conclusion Atasan Lyra memiliki nilai *Support* sebesar 0.11, dan *Confidence* 0.68, dan *Lift Ratio* sebesar 5.2. Gambar 9 menunjukkan hasil rekomendasi secara keseluruhan.



Gambar 3. Hasil Rekomendasi Produk

1. Implementasi Program

 Implementasi program menggunakan software VSCO dan XAMPP guna menerapkan rancangan yang sudah dibuat.

1. Halaman *Login*

Pada Gambar 4 Halaman Login, pengguna memasukkan username dan kata sandi untuk mengakses fitur aplikasi. Jika salah, akses ditolak. Fitur "Remember Me" mengingat username dan kata sandi untuk pengisian otomatis.



Gambar 4. Tampilan Halaman Login

1. Halaman *Dashboard*

Gambar 5 *Dashboard* adalah halaman utama yangmenampilkan beberapa informasi penting terkait performa penjualan toko Ceriastore seperti data penjualan, data clustering dan data assosiasi.



Gambar 5. Halaman Dashboard

1. Halaman Data Penjualan

Halaman ini digunakan untuk mengisi atau menginput data penjualan. Fitur-fiturnya termasuk "Import Data" untuk mengimpor data yang sudah ada dan "Tambah Data" untuk menambah data penjualan per transaksi.



Gambar 6. Tampilan Halaman Data Penjualan

1. Halaman *Clustering*

Pada Halaman *Clustering*, pengguna melakukan proses *clustering K-Medoids* pada data penjualan yang sudah disimpan. Pengguna dapat menentukan jumlah *cluster* yang diinginkan untuk memproses data penjualan.



Gambar 7. Tampilan Halaman Clustering

1. Halaman Asosiasi

Gambar 8 Halaman Asosiasi menghasilkan output dengan algoritma *FP-Growth*. Data yang telah di *cluster* otomatis masuk ke halaman ini, di mana pengguna dapat memasukkan *Confidence* dan *Minimum Support* untuk menemukan hasil yang diinginkan.



Gambar 8. Tampilan Halaman Asosiasi

1. Halaman Hasil

Pada Halaman Hasil, pengguna dapat melihat hasil yang telah di proses dari awal, halaman hasil menampilkan hasil clustering serta penjelasan dari proses yang telah ditentukan.dan hasil asosiasi produk yang lolos ditemukan nilai asosiasi itemset produk.



Gambar 9. Tampilan Halaman Hasil

# **IV. KESIMPULAN**

 Hasil penelitian menunjukkan bahwa pendekatan kombinasi K-Medoids dan FP-Growth efektif dalam mengelompokkan produk berdasarkan pola penjualan yang signifikan dan mengidentifikasi hubungan asosiasi antar produk yang relevan. Kluster penjualan (rendah, sedang, dan tinggi) memberikan wawasan penting untuk meningkatkan manajemen stok dan strategi pemasaran di Ceria Store. Rekomendasi produk yang dihasilkan dari analisis asosiasi dapat digunakan untuk mengarahkan strategi pemasaran yang lebih tepat sasaran, yang berpotensi meningkatkan penjualan dan kepuasan pelanggan.

 Saran yang dapat diberikan berupa pengelola tokok Ceria Store terus memperbarui analisis data secara berkala, mengoptimalkan stok produk berdasarkan temuan analisis, mengpersonalisasi strategi pemasaran, dan mengembangkan sistem teknologi untuk memperkuat keunggulan kompetitif mereka dalam industri ritel.

##### **DAFTAR PUSTAKA**

[1] A. Andirwan, V. Asmilita, M. Zhafran, A. Syaiful, and M. Beddu, “Strategi Pemasaran Digital: Inovasi untuk Maksimalkan Penjualan Produk Konsumen di Era Digital.”

[2] D. Transaksi Bongkar Muat di Provinsi Riau, I. Kamila, U. Khairunnisa, P. Studi Sistem Informasi, and F. Sains dan Teknologi UIN Sultan Syarif Kasim Riau, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan,” *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119–125, 2019.

[3] C. Eri Firman Akademi Manajemen Informatika and K. Dumai Jl Utama Karya Kel Bukit Batrem Kec Dumai Timur, “I N F O R M A T I K A PENENTUAN POLA YANG SERING MUNCUL UNTUK PENJUALAN PUPUK MENGGUNAKAN ALGORITMA FP-GROWTH,” *Jurnal Informatika, Manajemen dan Komputer*, vol. 9, no. 2, 2017.

[4] wiwit A. Triyanto, V. Suhartono, and H. Himawan, “Analisis Keranjang Pasar Menggna K-medoids dan FP-Growth,” *Pseudocode*, 2014.

[5] I. Syukra, A. Hidayat, and M. Z. Fauzi, “Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendations,” *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 2, no. 2, p. 107, Nov. 2019, doi: 10.24014/ijaidm.v2i2.8326.

[6] M. Imam Ghozali, fan N. Zaenal Ehwan SMA, J. Pati, and W. Harry Sugiharto, “ANALISA POLA BELANJA MENGGUNAKAN ALGORITMA FP GROWTH, SELF ORGANIZING MAP (SOM) DAN K MEDOIDS,” *Jurnal SIMETRIS*, vol. 8, 2017.

[7] W. A. Triyanto, “ALGORITMA K-MEDOIDS UNTUK PENENTUAN STRATEGI PEMASARAN PRODUK,” *Jurnal SIMETRIS*, vol. 6, 2015.

[8] S. Adibya, P. Putra, P. Kasih, and J. Sahertian, “Implementasi Pola Penjualan Barang di Minimarket Menggunakkan Metode Apriori.”

[9] I. W. Septiani, Abd. C. Fauzan, and M. M. Huda, “Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, no. 4, p. 556, Jul. 2022, doi: 10.30865/json.v3i4.4055.

 [10] A. R. Wibowo and A. Jananto, “IMPLEMENTASI DATA MINING METODE ASOSIASI ALGORITMA FP-GROWTH PADA PERUSAHAAN RITEL,” vol. 10, no. 2, pp. 200–212, 2020.