

Implementasi Data Mining Analisa Pola Belanja Customer Dengan menggunakan FP-Growth pada Produk Fashion

Ririn Agustin¹, Adam Arif Budiman^{2*}

^{1,2} Jurusan Teknologi Informasi, Universitas Darma Persada
^{1,2} Jl. Taman Malaka Selatan, Jakarta Timur, DKI Jakarta 13450, Indonesia
*ariadam@gmail.com

Abstrak : Penelitian ini menerapkan data mining untuk menganalisis pola pelanggan dan prediksi produk fashion. Metode FP-Growth digunakan untuk mengidentifikasi pola itemset yang sering muncul, dataset berisi riwayat pembelian pelanggan dan atribut produk fashion. Hasil analisis pola pelanggan dan prediksi produk fashion dapat membantu perusahaan fashion dalam mengambil keputusan strategis. Penelitian ini memberikan kontribusi pada penggunaan data mining untuk memahami preferensi pelanggan dan meningkatkan keputusan bisnis bagi perusahaan fashion. Penggunaan dataset yang terdiri dari riwayat pembelian pelanggan dan atribut produk fashion. Pertama, dengan menggunakan algoritma FP-Growth, dilakukan analisis untuk mengidentifikasi pola itemset yang sering muncul pada data pelanggan. Hasil analisis digunakan untuk memahami preferensi dan kebiasaan belanja pelanggan.

Kata Kunci – Data Mining, Analisis Pola, FP-Growth

Copyright © 2018 JURNAL TIFDA
All rights reserved.

I. PENDAHULUAN

Fashion adalah gaya berpakaian yang digunakan seseorang setiap hari untuk meningkatkan penampilannya, baik di acara tertentu maupun di kehidupan sehari-hari. Fashion memiliki jangkauan luas tentang mode untuk semua usia. Seiring berjalannya waktu, mode berkembang. Pecinta fashion bereksresi dengan selera pakaian yang berbeda-beda untuk setiap musim karena berbagai era yang terjadi setiap tahun.

Berbagai gaya pakaian menjadi identitas penggunaannya. Orang sering menggunakan gaya casual, *vintage*, formal, dan *sporty*. Gaya berpakaian seseorang dapat dipengaruhi oleh banyak faktor, termasuk preferensi musik, media massa, dan internet, serta budaya. Memilih produk berdasarkan prioritas juga merupakan faktor lain. Karena preferensi pengguna yang beragam, hal ini membuat keadaan pasar tidak stabil. Keberagaman fashion memungkinkan perusahaan memiliki merek untuk menarik pembeli.

Choice Fashion memiliki banyak pertimbangan tentang produk yang akan dijual di pasar. Sebagai reseller dan distributor, mereka menawarkan banyak koleksi produk dari berbagai merek fashion. Seleksi merek tanpa acuan akan memakan waktu dan produk yang tersedia tidak akan memenuhi target pasar. Memilih produk untuk dijual dan dipromosikan dapat

dibantu dengan kriteria seperti harga, kegunaan, bahan, dan gender. dengan Algoritma data mining dapat dibuat suatu aplikasi berbasis web untuk menentukan merek yang paling populer. Pada penelitian ini digunakan algoritma FP-Growth untuk menentukan pola penjualan fashion.

II. METODOLOGI

A. Data mining

Data mining adalah proses menemukan pola-pola yang menarik dan tersembunyi (pola tersembunyi) dalam sejumlah besar data yang disimpan dalam basis data, data warehouse, atau tempat penyimpanan data lainnya [1].

Menurut Sumanthi dan Sivandham (2009), data mining juga didefinisikan sebagai bagian dari proses penggalian pengetahuan dalam database. Proses KDD menggabungkan berbagai metode, seperti statistik, basis data, kecerdasan buatan (AI), pembelajaran mesin, pengenalan pola (Pattern Recognition), pemodelan yang menangani ketidakpastian, visualisasi data, dan kecerdasan buatan (AI).

Secara sederhana, data mining adalah proses ekstraksi untuk mendapatkan informasi penting yang implisit dan belum diketahui. Data mining dimulai

dengan proses pemilihan, pembersihan, pre-processing, dan transformasi data. Semua ini merupakan bagian dari proses yang ada dalam KDD [2].

CRISP-DM (Cross Industry Standard Process for Data Mining) dibuat pada tahun 1996 berdasarkan analisis dari berbagai industri, seperti SPSS, NCR, dan standarisasi Daimler Chrysler (Daimler-Benz). CRISP-DM menawarkan strategi pemecahan masalah umum untuk bisnis atau unit penelitian [3].

B. Algoritma FP-Growth

FP-Growth algoritma data mining digunakan untuk menemukan pola asosiasi dalam dataset [4], terutama dalam analisis keranjang belanja. Ini adalah pengembangan dari algoritma Apriori dan dimaksudkan untuk memperbaiki kekurangannya, terutama terkait efisiensi waktu [5]. FP-Growth menggunakan FP-Tree (*Frequent Pattern Tree*) sebagai struktur data untuk menyimpan informasi tentang itemsets yang sering muncul tanpa perlu melakukan pemindaian database berulang kali [6]. Banyak industri menggunakan algoritma ini, seperti analisis penjualan untuk mengidentifikasi pola pembelian konsumen; misalnya, FP-Growth berhasil menemukan aturan asosiasi antara produk yang sering dibeli bersamaan [7]. Algoritma ini dapat membantu dalam pengelolaan stok dan strategi pemasarannya.

III. HASIL DAN PEMBAHASAN

Dalam melakukan penelitian pada penerapan dengan menggunakan Algoritma Frequent Pattern Growth beberapa tahap yang dilakukan. Pada tahap pertama yaitu menyiapkan sebuah sampel dataset dengan mengambil data transaksi penjualan yang ada di toko Fashion yang bernama Choice Fashion. Data yang di dapatkan ada 23 data brand produk, sedangkan data pada transaksi penjualan 335 yang diambil 1 brand setiap bulannya untuk data sampel. Data yang dibutuhkan untuk dataset yang berisikan 23 data brand produk yang dijual oleh Choice Fashion. Berikut ini dapat dilihat pada Tabel 1.

Pada Tabel 2 menunjukkan bahwa adalah data sampel dataset transaksi penjualan 23 brand produk Choice Fashion. Setelah terdapat data transaksi yang sudah di ketahui, maka selanjutnya untuk menentukan kemunculan daari data setiap itemset yang ada dengan nilai support lebih dari 35% dari 335 data transaksi tersebut. Berikut ini dapat dilihat pada Tabel 3.

Tabel 1. nama *brand*

No	Kode	Nama Brand
1	A01	Adidas
2	A02	Airwalk
3	A03	Cardinal
4	A04	Converse
5	A05	Cotton On
6	A06	Dc
7	A07	Diadora
8	A08	Expand
9	A09	Hush Puppies
10	A10	Logo Jeans
11	A11	Manggo
12	A12	Mark & Spancer
13	A13	New Balance
14	A14	Nike
15	A15	Ninety Degrees
16	A16	Pierox
17	A17	Quiksilver
18	A18	Reebok
19	A19	Seven Level
20	A20	Skechers
21	A21	Superga
22	A22	Tolliver
23	A23	Vans

Tabel 2. Sampel Dataset transaksi penjualan

Tanggal	Brand Terjual
Tanggal 1-16 Jan	A02, A04, A09, A11, A12, A22, A23
Tanggal 17-31 Jan	A01, A09, A14, A23
Tanggal 1-14 Feb	A04, A05, A09, A14, A22
Tanggal 15-28 Jan	A01, A23
Tanggal 1-16 Maret	A01, A03, A04, A06, A09, A13, A14, A15
Tangga 17-31Mar	A01, A09, A10, A12, A14, A17, A22, A23
Tanggal 1-15 April	A01, A02, A03, A04, A07, A09, A10, A13, A14,A15, A17, A18, A20, A21, A22, A23
Tanggal 16-31 April	A01, A02, A03, A04, A08, A09, A14, A18, A20, A23
Tanggal 1-16 Mei	A01, A03, A04, A05, A07, A09, A13, A14, A20, A22, A23
Tanggal 17-31 Mei	A03, A04, A05, A07, A09, A14, A19, A20
Tanggal 1-15 Jun	A02, A04, A05, A07, A09, A12, A14, A15, A16, A20, A22, A23
Tanggal 16-30 Jun	A02, A04, A09, A14, A15, A20, A23

Setelah melakukan tahap pada perhitungan mencari frekuensi kemunculan setiap itemset telah di ketahui produk dengan nilai *support count* > 35% dimana adai tem yang tidak memenuhi syarat nilai yaitu A06, A08, A11, A16, A19, A21, A10, A17, A18, A12, A13, A05, A07, A05, A07, A15. Maka dari setiap item yang memenuhi syarat dapat dilakukan tahap pembuatan FP-Tree.

Tabel 3 Frekuensi itemset dengan nilai support > 35%.

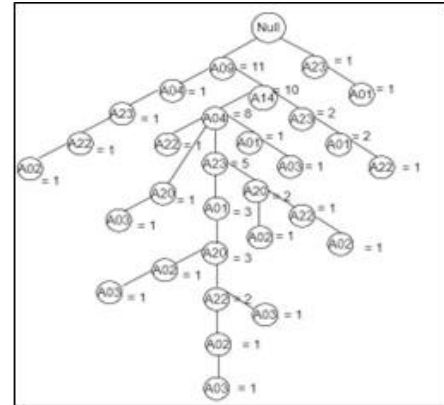
Brand	Frekuensi	Proses Support	Support
A01	7	(7/12)*100%	58,33%
A02	5	(5/12)*100%	41,67%
A03	5	(5/12)*100%	41,67%
A04	9	(9/12)*100%	75,00%
A05	4	(4/12)*100%	33,33%
A06	1	(1/12)*100%	8,33%
A07	4	(4/12)*100%	33,33%
A08	1	(1/12)*100%	8,33%
A09	11	(11/12)*100%	91,67%
A10	2	(2/12)*100%	16,67%
A11	1	(1/12)*100%	8,33%
A12	3	(3/12)*100%	25,00%
A13	3	(3/12)*100%	25,00%
A14	10	(10/12)*100%	83,33%
A15	4	(4/12)*100%	33,33%
A16	1	(1/12)*100%	8,33%
A17	2	(2/12)*100%	16,67%
A18	2	(2/12)*100%	16,67%
A19	1	(1/12)*100%	8,33%
A20	6	(6/12)*100%	50,00%
A21	1	(1/12)*100%	8,33%
A22	6	(6/12)*100%	50,00%
A23	9	(9/12)*100%	75,00%

Tahapan selanjutnya adalah mendata kemunculan pada item berdasarkan frequent tertinggi, setelah dilakukan pengurutan item dengan nilai support > 35%, berikut dapat dilihat pada tabel 4

Tabel 4. Pengurutan frequent tertinggi

Brand	Frekuensi	Proses Support	Support
A09	11	(11/12)*100%	91,67%
A14	10	(10/12)*100%	83,33%
A04	9	(9/12)*100%	75,00%
A23	9	(9/12)*100%	75,00%
A01	7	(7/12)*100%	58,33%
A20	6	(6/12)*100%	50,00%
A22	6	(6/12)*100%	50,00%
A02	5	(5/12)*100%	41,67%

Tabel 4 diatas menunjukkan bahwa pencarian data item telah menghitung dengan nilai support dari setiap item dan memenuhi syarat minimum tersebut. Maka hasil pencarian frequent item set sudah didapatkan frekuensi kemunculan dari setiap item mengurut yang terbesar ke terkecil. Setelah melakukan proses pada pencarian dataset item transaksi dengan syarat minimum support, maka tahap yang dilakukan selanjutnya adalah membangun Frequent Pattern Tree. Pada Gambar 1 menunjukan Pembentukan FP-Tree dan setelah melakukan pembentukan FP-Tree, maka langkah selanjutnya adalah pembangkitan *conditional pattern* yang dapat dilihat pada Tabel 5



Gambar 1. FP-Tree

Tabel 5. *Conditional pattern*

Brand	Condition Pattern Base
A14	(A09 = 10)
A04	(A09 = 1) (A14, A09 = 8)
A23	(A04, A09 = 1) (A04, A14, A09 = 5) (A14, A09 = 2)
A01	(A23, A04, A14, A09 = 3) (A04, A014, A09=1) (A23, A14, A09=2) (A23 = 1)
A20	(A04, A14, A09 = 1) (A01, A23, A04, A14, A09 = 3) (A23, A04, A14, A09 = 2)
A22	(A23, A04, A09 = 1) (A04, A14, A09 = 1) (A20, A01, A23, A04, A14, A09 = 2) (A20, A23, A04, A014, A09 = 1) (A01, A23, A14, A09 = 1)
A02	(A22, A23, A04, A09 = 1) (A20, A01, A23, A04, A14, A09 = 1) (A22, A20, A01, A23, A04, A14, A09 = 1) (A20, A23, A04, A14, A09 = 1) (A22, A20, A23, A04, A14, A09 = 1)

Tahap pembangkitan conditional FP-Tree Setelah melakukan pencarian *Conditional Pattern Base*, maka selanjutnya dilakukan untuk menjumlahkan dari setiap item yang memiliki *support count* yang ada dengan nilai support > 35% akan di bangkitkan dengan FP-Tree. Berikut ini dapat dilihat di Tabel 6

Tabel 6 Tabel Hasil Pembangkitan Conditional FP-Tree

Brand	Condition Pattern FP-Tree
A14	(A14.A09 = 10)
A04	(A09 = 1) (A14, A09 = 8)
A23	(A04, A09 = 1) (A04, A14, A09 = 5) (A14, A09 = 2)
A01	(A23, A04, A14, A09 = 3) (A04, A014, A09=1) (A23, A14, A09=2) (A23 = 1)
A20	(A04, A14, A09 = 1) (A01, A23, A04, A14, A09 = 3) (A23, A04, A14, A09 = 2)
A22	(A23, A04, A09 = 1) (A04, A14, A09 = 1) (A20, A01, A23, A04, A14, A09 = 2) (A20, A23, A04, A014, A09 = 1) (A01, A23, A14, A09 = 1)
A02	(A22, A23, A04, A09 = 1) (A20, A01, A23, A04, A14, A09 = 1) (A22, A20, A01, A23, A04, A14, A09 = 1) (A20, A23, A04, A14, A09 = 1) (A22, A20, A23, A04, A14, A12, A09 = 1)

Tahap selanjutnya yang dilakukan adalah menentukan atau mencari frekuensi antar item. Dimana tahap ini mencari *single path* dikombinasikan atau direlasikan dengan item yang sudah diketahui sebelumnya *Conditional FP-Tree*. Berikut ini dapat dilihat pada Table 7

Tabel 7. *Frequent Itemset Pattern Generated*

Brand	<i>Frequent Itemset Pattern Generate</i>
A14	(A14, A09=7)
A04	(A04, A09 = 9) (A04, A14 = 8)
A23	(A23, A04 = 6) (A23, A09 = 8) (A23, A14 = 7)
A01	(A01, A23 = 6)(A01, A09=6)
A20	(A20, A04 = 6) (A20, A14 = 6) (A20, A09 = 6) (A20, A23 = 5)
A22	(A22, A09 = 6)
A02	(A02, A23 = 5) (A02, A04 = 5) (A02, A09 = 5)

Tahapan ini dilakukan untuk mencari *single path* dikombinasikan dengan itemset yang sudah didapatkan dari *conditional pattern FP-Tree*. Berikut daftar item yang memenuhi syarat tersebut dituliskan pada Table 8.

Tabel 8. hasil *association Rule*

Item Set	P(A B)	Support	A	Process Confidence	Confidence
(A04, A09)	9	75,00%	9	(9/9)*100%	100,00%
(A20, A04)	6	50,00%	6	(6/6)*100%	100,00%
(A20, A14)	6	50,00%	6	(6/6)*100%	100,00%
(A20, A09)	6	50,00%	6	(6/6)*100%	100,00%
(A22, A09)	6	50,00%	6	(6/6)*100%	100,00%
(A02, A23)	5	41,67%	5	(5/5)*100%	100,00%
(A02, A04)	5	41,67%	5	(5/5)*100%	100,00%
(A02, A09)	5	41,67%	5	(5/5)*100%	100,00%
(A04, A14)	8	66,67%	9	(8/9)*100%	88,89%

Pada Table 8 hasil assosiasi yang memiliki nilai support > 35% dan nilai *confidence* > 75% yang memenuhi syarat berjumlah 9 rules dan 2 rules dengan nilai yang tidak memenuhi syarat.

IV. KESIMPULAN

Penelitian tentang Analisa Pola Belanja Customer dengan menggunakan FP-Growth pada Produk Fashion ini dirancang untuk memberikan strategi promosi dan memperkuat *branding*. Dengan data mining dan analisis pola pembelian, sistem memberikan wawasan tentang tren dan preferensi pelanggan untuk pengembangan strategi promosi yang efektif. Disamping itu sistem ini mempermudah *branding* produk Fashion berdasarkan pola pembelian dengan menganalisis pola pembelian, sistem mengidentifikasi kombinasi produk populer dan hubungan antarproduk yang berfokus kepada strategi branding dan penempatan produk dengan lebih tepat.

DAFTAR PUSTAKA

- [1] Tan, dkk (2006). Introduction to Data Mining. Pearson Education, Inc.
- [2] Sumathi, S dkk (2006). Introduction to Data Mining and its Applications. Spinger, Verlag Berlin Heidelberg.
- [3] Larose, Daniel (2006). Discovering Knowledge in Data: An Introduction to Data mining: Jhon Wiley & Son, Inc., USA
- [4] Lintang dkk, Penerapan Algoritma FP-Growth Untuk Menentukan Pola Penjualan Toko Ellia Umami, *Journal of Student Research (JSR)*, Vol.1, No.3 Mei 2023, e-ISSN: 2963-9697; p-ISSN: 2963-9859, Hal 367-378
- [5] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, no. 1, pp. 34–39, Jan. 1959.
- [6] E. P. Wigner, "Theory of traveling-wave optical laser," *Phys.Rev.*, vol. 134, pp. A635–A646, Dec. 1965.
- [7] E. H. Miller, "A note on reflector arrays," *IEEE Trans. Antennas Propagat.*, to be published.