

Klasterisasi Data Penjualan Barang Menggunakan Algoritma K-Medoids Untuk Mengoptimalkan Penjualan Sepatu

Brigitto Julio Niceforus¹, Fadli Hafizh Sidiq², Wisnu Wicaksono³, Tedy Setiadi⁴

^{1,2,3}Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan, Jln. Ring Road
Selatan, Tamanan Banguntapan Bantul Yogyakarta, 55166

tedy.setiady@tif.uad.ac.id

Abstrak

Persediaan barang menumpuk karena kesulitan manajemen mengambil kebijakan untuk menyuplai barang. Penumpukan barang menyebabkan harga barang turun. Oleh karena itu, teknik yang diperlukan untuk melakukan analisis penjualan barang yang berguna untuk menyelesaikan masalah saat ini. Analisis penjualan barang adalah tujuan dari penelitian ini. Algoritma K-Medoids akan mengelompokkan data penjualan untuk membantu teknik analisis menganalisis penjualan barang. Analisis variabel yang dilakukan membagi kelompok data menjadi dua: barang yang laris dan barang yang tidak laris. Penelitian ini menggunakan 30 sampel data penjualan barang. Hasilnya menawarkan informasi tentang barang yang laris dan barang yang tidak laris, yang membantu menyelesaikan masalah dengan ketersediaan stok barang.

Kata kunci: Algoritma K-Medoids; Penjualan barang; Stok Barang

A. Pendahuluan

Data Mining merupakan teknik yang menggunakan teknik kecerdasan buatan, *statistic*, matematika dan *Machine Learning*, untuk menggali dan mengidentifikasi informasi yang berharga dan bermanfaat serta informasi-informasi penting yang ada dalam data yang besar. Bisnis dapat menggunakan teknologi informasi dalam banyak bidang, seperti dalam pemasaran produk dan barang, yang menghasilkan banyak data. Pelaku bisnis harus mahir dalam pengolahan data penjualan untuk mengoptimalkan penjualan barang. Clustering adalah salah satu metode data mining yang paling populer. Dalam statistik dan mesin pembelajaran, pengelompokan K-Means merupakan metode analisis kelompok yang mengarah pada pemartisian N objek pengamatan ke dalam K kelompok (cluster) dimana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat, mirip dengan algoritma Expectation-Maximization untuk Gaussian Mixture di mana keduanya mencoba untuk menemukan pusat dari kelompok dalam data sebanyak iterasi perbaikan yang dilakukan oleh kedua algoritma.

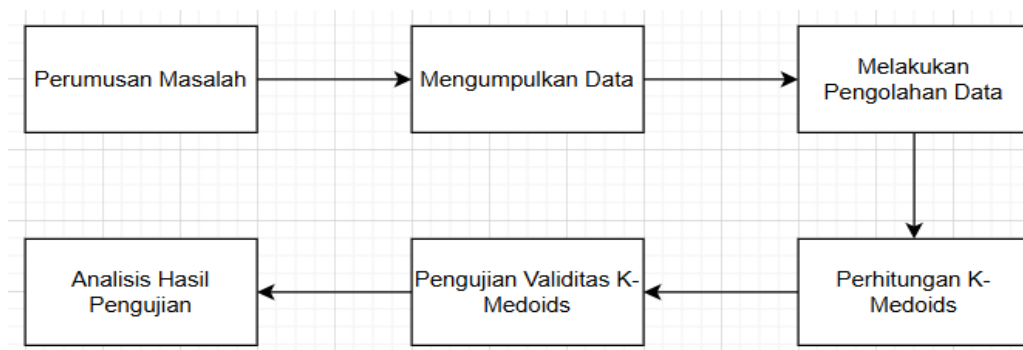
Salah satu masalah utama dengan clustering adalah ketidakmampuan untuk menentukan jumlah cluster yang tepat berdasarkan data yang diuji. Data mining adalah proses yang dapat secara otomatis menganalisis dan mengekstrak data menggunakan teknik pembelajaran komputer (*machine learning*). Algoritma K-Medoids atau sering disebut juga dengan algoritma PAM (*Partitioning Around Medoid*) dikembangkan oleh Leonard Kaufman dan Peter J.

Rousseeuw yang dimana algoritma ini mirip dengan *k-means*, terutama keduanya merupakan algoritma partitional dengan kata lain, keduanya memecah dataset menjadi kelompok-kelompok.

K-Medoids merupakan algoritma yang mempunyai strategi dasar untuk menemukan k cluster dalam n objek dengan terlebih dahulu menemukan perwakilan secara acak medoids pada setiap cluster. Oleh karena itu, prinsip yang mendasari metode partisi masih dapat diterapkan, yaitu meminimalkan jumlah perbedaan antara setiap objek dan titik referensi yang sesuai (medoid). Penelitian ini bertujuan untuk bagaimana mengelompokkan data penjualan dengan algoritma kmedois. Dengan menganalisis data penjualan, maka akan dapat ditentukan mana yang perlu diperbaiki dalam hal penjualan. Informasi ini kemudian dapat digunakan oleh pihak e-commers sebagai acuan dalam menyusun strategi pemasaran yang bertujuan untuk meningkatkan produk tersebut dalam penjualan[4]. Tujuan dari pengelompokkan sekumpulan data objek kedalam beberapa kelompok yang mempunyai karakteristik tertentu dan dapat dibedakan satu sama lainnya adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan.

B. Metode

Metode penelitian akan dijelaskan dalam kerangka penelitian, yang terdiri dari urutan langkah- langkah dalam mengambil bagian dalam penelitian. Image 1 menunjukkan struktur penelitian ini.



Gambar 1 struktur penelitian

Gambar 1 menunjukkan struktur penelitian, yang diuraikan sebagai berikut, seperti menciptakan masalah, mengumpulkan data, mengolah data, melakukan perhitungan K-Medoids, menguji validitas K-Medoids, dan menganalisis hasil pengujian. Langkah-langkah penelitian ini adalah sebagai berikut:

1. Perumusan Masalah

Pada tahap ini, masalah diidentifikasi melalui analisis. Langkah ini dilakukan untuk mendapatkan pemahaman yang lebih baik tentang subjek penelitian dan menentukan apakah subjek tersebut sesuai dengan topik penelitian yang dipilih peneliti.

2. Mengumpulkan Data

Pada tahap ini, data akan dikumpulkan untuk mendukung penelitian ini. Data primer dan sekunder terdiri dari data yang dikumpulkan. Informasi primer dapat diperoleh melalui pengamatan langsung di lapangan.

3. Melakukan Pengolahan Data

Untuk menyelesaikan tahap ini, data diproses terlebih dahulu. Tahap berikutnya dalam penambangan data dikenal sebagai pemrosesan data. Tujuan dari tahap ini adalah untuk meningkatkan kualitas data yang terkumpul melalui penggunaan algoritma K-Medoids. Proses tersebut mencakup pemilihan fitur, pembersihan data dan pengolahan data.

4. Perhitungan K-Medoids

Perhitungan algoritma K-Medoids dilakukan menggunakan rumus yang telah ada sebelumnya. Algoritma K-Medoids membagi data secara sistematis ke dalam cluster. Setelah cluster terbentuk, pola clustering dianalisis, dan informasi yang diperoleh dari analisis ini digunakan untuk membuat kebijakan analisis penjualan.

5. Pengujian Validitas K-Medoids

Pada tahap ini, data sebelumnya digunakan untuk melakukan pengujian kelompok. Berdasarkan temuan ini, pengujian, kesimpulan dibuat melalui penghitungan algoritma L-Medoids.

6. Analisis Hasil Pengujian

Pada tahap analisis klaster, klaster yang paling dicari dijelaskan berdasarkan karakteristik yang terbentuk, pola untuk mengetahui hubungan yang terbentuk dari hasil klaster, analisis dilakukan.

C. Hasil dan Pembahasan

Pada tahap awal, data penjualan sepatu dikelompokkan atau disegmentasi pada Konveksi. Liberty Shoemakers adalah nilai dari faktor stok dan penjualan. Gambar 2 menunjukkan data sampel yang digunakan untuk percobaan pencacahan manual.

No	Nama Produk	Terjual	Stok
1	Adidas Running Adizero	30	100
2	Adidas Running Duramo	50	120
3	Adidas Running Supernova 3	30	90
4	Adidas Lifestyle GRAND COURT 2.0	10	100
5	Adidas Lifestyle NY 90	20	80
6	ASICS Gel-Kayano 31 Standar	25	95
7	ASICS Gel-Cumulus 26 Standar	45	150
8	ASICS Gel-Kinsei Max Standar	15	85
9	ASICS Gel-Nimbus 26 Standar	25	150
10	ASICS Gel-Quantum 90	18	50
11	Reebok Club C 85 Vintage	15	55
12	Reebok Classic Leather	16	60
13	Reebok Court Advance	21	80
14	Reebok Smash Edge S	20	110
15	Reebok Classic Nylon	12	90
16	Skechers Uno	21	70
17	Skechers Cordova Classic	25	80
18	Skechers Go Run Consistent 2.0	21	100
19	Skechers Skech-Air Ventura	15	80
20	Skechers Vapor Lite	25	175
21	Skechers D'LITES 4.0	30	100

Gambar 2. Contoh Data Sampel

Langkah berikutnya, transformasi data, adalah bagian dari proses transformasi data. Ini berarti mengubah format data dari sumbernya untuk memenuhi persyaratan pemrosesan metode yang digunakan. Dalam penelitian ini, beberapa teknik transformasi data dinormalisasi dengan persamaan min-max. Normalisasi data mengurangi batas bawah 0 dan batas atas 1. Hasil normalisasi data disajikan pada gambar 3.

Nama Produk	Nilai	
	Terjual	Stok
Adidas Running Adizero	0.5	0.4
Adidas Running Duramo	1	0.56
Adidas Running Supernova 3	0.5	0.32
Adidas Lifestyle GRAND COURT 2.0	0	0.4
Adidas Lifestyle NY 90	0.25	0.24
ASICS Gel-Kayano 31 Standar	375	0.36
ASICS Gel-Cumulus 26 Standar	875	0.8
ASICS Gel-Kinsei Max Standar	125	0.28
ASICS Gel-Nimbus 26 Standar	375	0.8
ASICS Gel-Quantum 90	0.2	0
Reebok Club C 85 Vintage	125	0.04
Reebok Classic Leather	0.15	0.08
Reebok Court Advance	275	0.24
Reebok Smash Edge S	0.25	0.48
Reebok Classic Nylon	0.05	0.32
Skechers Uno	275	0.16
Skechers Cordova Classic	375	0.24
Skechers Go Run Consistent 2.0	275	0.4
Skechers Skech-Air Ventura	125	0.24
Skechers Vapor Lite	375	1
Skechers D'LITES 4.0	0.5	0.4

Gambar 3. Data Hasil Normalisasi

a. Penentuan pusat awal cluster

Penentuan pusat awal cluster dilakukan pengambilan secara acak dari data Tabel 2, yaitu diasumsikan:

Pusat cluster 1: (0.5, 0.4)

Pusat cluster 2: (1, 0.56)

b. Perhitungan euclidean distance

Euclidean distance digunakan untuk mengukur jarak antara data dengan pusat cluster. Rumus Euclidian Distance sebagai berikut (Hadi & Diana, 2020):

$$De = \sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2} \quad (1)$$

dimana De adalah Euclidean Distance, i adalah banyaknya objek, (x,y) adalah koordinat objek (s,t). Dari 21 data yang dijadikan sampel telah dipilih pusat awal cluster yaitu C1= (0.5, 0.4), C2=(1, 0.56). Kemudian menghitung jarak dari sisa data sampel ke pusat cluster. Sehingga diperoleh tabel jarak dari Medoid. Jarak dari Medoid disajikan pada gambar 4.

Nama Produk	Terjual	Stok	Kedekatan	Cluster
Adidas Running	0, 0000	0, 5250	0, 0000	1
Adidas Running	0, 5250	0, 0000	0, 0000	2
Adidas Running	0, 0800	0, 5546	0, 0800	1
Adidas Lifestyle	0, 5000	1, 0127	0, 5000	1
Adidas Lifestyle	0, 2968	0, 8154	0, 2968	1
ASICS Gel-Kayan	0, 1312	0, 6562	0, 1312	1
ASICS Gel-Cumu	0, 5483	0, 2706	0, 2706	2
ASICS Gel-Kinsei	0, 3937	0, 9187	0, 3937	1
ASICS Gel-Nimbu	0, 4191	0, 6695	0, 4191	1
ASICS Gel-Quant	0, 5000	0, 9765	0, 5000	1
Reebok Club C 8	0, 5198	1, 0179	0, 5198	1
Reebok Classic L	0, 4742	0, 9762	0, 4742	1
Reebok Court Ac	0, 2761	0, 7925	0, 2761	1
Reebok Smash E	0, 2625	0, 7543	0, 2625	1
Reebok Classic N	0, 4571	0, 9798	0, 4571	1
Skechers Uno	0, 3290	0, 8280	0, 3290	1
Skechers Cordov	0, 2030	0, 7022	0, 2030	1
Skechers Go Run	0, 2250	0, 7424	0, 2250	1
Skechers Skech-A	0, 4077	0, 9317	0, 4077	1
Skechers Vapor I	0, 6129	0, 7643	0, 6129	1
Skechers D'LITES	0, 0000	0, 5250	0, 0000	1

Gambar 4. Hasil Jarak Medoid

Gambar 4. menunjukkan kolom kedekatan nilai diperoleh dari nilai min (cost 1, cost 2). Sedangkan untuk kolom Cluster nilai diperoleh dari nilai kolom keterangan yang terdapat pada kolom cost. Jika nilai keterangan terdapat pada cost 1 maka nilai clustering bernilai 1. Jika nilai keterangan terdapat pada cost 2 maka nilai Clustering bernilai 2. Dari hasil clustering maka diperoleh hasil 19 record pada cluster 1, dan 2 record pada cluster 2. Pada ketentuan Tabel 3, cluster 1 itu sama dengan Tidak Laris dan cluster 2 sama dengan Laris dan dapat disajikan pada gambar 5.

Nama Produk	Terjual	Stok	Kedekatan	Cluster
Adidas Running	0, 0000	0, 5250	0, 0000	Kurang Minat
Adidas Running	0, 5250	0, 0000	0, 0000	Minat
Adidas Running	0, 0800	0, 5546	0, 0800	Kurang Minat
Adidas Lifestyle	0, 5000	1, 0127	0, 5000	Kurang Minat
Adidas Lifestyle	0, 2968	0, 8154	0, 2968	Kurang Minat
ASICS Gel-Kayan	0, 1312	0, 6562	0, 1312	Kurang Minat
ASICS Gel-Cumu	0, 5483	0, 2706	0, 2706	Minat
ASICS Gel-Kinsei	0, 3937	0, 9187	0, 3937	Kurang Minat
ASICS Gel-Nimbu	0, 4191	0, 6695	0, 4191	Kurang Minat
ASICS Gel-Quant	0, 5000	0, 9765	0, 5000	Kurang Minat
Reebok Club C 8	0, 5198	1, 0179	0, 5198	Kurang Minat
Reebok Classic L	0, 4742	0, 9762	0, 4742	Kurang Minat
Reebok Court Ac	0, 2761	0, 7925	0, 2761	Kurang Minat
Reebok Smash E	0, 2625	0, 7543	0, 2625	Kurang Minat
Reebok Classic N	0, 4571	0, 9798	0, 4571	Kurang Minat
Skechers Uno	0, 3290	0, 8280	0, 3290	Kurang Minat
Skechers Cordov	0, 2030	0, 7022	0, 2030	Kurang Minat
Skechers Go Run	0, 2250	0, 7424	0, 2250	Kurang Minat
Skechers Skech-A	0, 4077	0, 9317	0, 4077	Kurang Minat
Skechers Vapor I	0, 6129	0, 7643	0, 6129	Kurang Minat
Skechers D'LITES	0, 0000	0, 5250	0, 0000	Kurang Minat

Gambar 5. Analisis Data

D. Kesimpulan

Penelitian ini berhasil menunjukkan bahwa penggunaan algoritma K-Medoids dalam klusterisasi data penjualan sepatu dapat mengoptimalkan strategi pemasaran dan pengelolaan persediaan. Dengan mengelompokkan data penjualan berdasarkan karakteristik pelanggan dan pola pembelian, algoritma K-Medoids mampu mengidentifikasi segmen pasar yang lebih tepat dan lebih terfokus. Kluster-kluster yang terbentuk memberikan wawasan yang mendalam mengenai preferensi pelanggan terhadap jenis sepatu tertentu, yang memungkinkan perusahaan untuk menyesuaikan promosi dan stok produk sesuai dengan kebutuhan pasar.

Daftar Pustaka

- B. Arief Setiawan, *PERBANDINGAN CLUSTERING OPTIMALISASI STOK BARANG MENGGUNAKAN ALGORITMA K-MEANS DAN ALGORTIMA K-MEDOIDS (STUDI KASUS : KLINIK BEN WARAS)*.
- S. Syahputra, S. Ramadani, A. Manaor, and H. Pardede, "MENENTUKAN STRATEGI PROMOSI MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS," *JOISIE Journal Of Information System And Informatics Engineering*, vol. 4, no. 1, pp. 7-14, 2020. "KOMPARASI ANALISIS K-MEDOIDS CLUSTERING."
- R. Narasati, R. Lestari, R. Herdiana, R. Hamonangan, and S. Anawar, "PENERAPAN ALGORITMA K-MEDOID PADA PENJUALAN KERUDUNG E-COMMERCE SHOPEE: FIFAU HIJAB," 2024.
- R. U. Br. Barus, I. Gunawan, B. E. Damanik, I. Parlina, and W. Saputra, "Pengelompokan Data Penjualan Mie Berdasarkan Bulan Dengan Menggunakan Algoritma K-Medoids," *Jurnal Ilmu Komputer dan Informatika*, vol. 1, no. 2, pp. 141-156, Nov. 2021, doi: 10.54082/jiki.15.