

## Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering

Hani Prastiwi<sup>1</sup>, Jeny Pricilia<sup>2</sup>, Errissya Raswir<sup>3</sup>

<sup>123</sup>Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Kota Jambi, indonesiaEmail : [hany89276@gmail.com](mailto:hany89276@gmail.com) [jenypricilia737@gmail.com](mailto:jenypricilia737@gmail.com), [errissya.raswir@gmail.com](mailto:errissya.raswir@gmail.com)

Email Penulis Korespondensi : [hany89276@gmail.com](mailto:hany89276@gmail.com)

**Abstrak-** Data mining adalah metode untuk menemukan pola tertentu dari kumpulan data yang berjumlah besar. Meskipun banyak dipelajari pada bidang ilmu komputer dan statistika, data mining adalah metode yang bisa diterapkan dan mempermudah pekerjaan di bidang lainnya juga. Namun sebenarnya apa itu data mining? Data mining adalah metode dalam ilmu komputer yang biasa digunakan dalam proses pencarian *knowledge*. Tahapan di dalamnya berguna untuk mencari pola-pola tertentu dari data yang ada pada *database*. Biasanya, metode ini banyak ditemukan pada bidang *machine learning* dan statistika. Metode K-Means sudah menjadi salah satu teknik data mining yang digunakan untuk merancang strategi persediaan barang yang efektif menggunakan data transaksi penjualan. Tujuan dari penelitian ini adalah untuk menerapkan algoritma K-Means dan data transaksi penjualan dari MM Glory sebagai contoh tipikal. Hasil analisis untuk penelitian ini menggunakan dua puluh (20) buah data.

Kata kunci: Algoritma, Data mining, K-Means, Stok

**Abstract-** Data mining is a method for finding certain patterns from large data sets. Although much studied in the fields of computer science and statistics, data mining is a method that can be applied and facilitates work in other fields as well. But what exactly is data mining? Data mining is a method in computer science that is commonly used in the knowledge search process. The stages in it are useful for finding certain patterns from the data in the database. Usually, this method is found in the fields of machine learning and statistics. The K-Means method has become one of the data mining techniques used to design an effective inventory strategy using sales transaction data. The purpose of this study is to apply the K-Means algorithm and sales transaction data from MM Glory as typical examples. The results of the analysis for this study used twenty (20) pieces of data.

Keywords: Algoritma, Data mining, K-Means, Stoc

### 1. PENDAHULUAN

Minimarket adalah swalayan atau toko kecil yang menyediakan produk barang atau jasa yang dibutuhkan dalam kehidupan sehari-hari oleh masyarakat. Definisi ini serupa dengan definisi dari toko kelontong, hanya saja mini market adalah bentuk modern dari toko kelontong. Pada kebanyakan minimarket yang ada di Jambi saat ini belum mempunyai metode baku yang diterapkan. Pemeriksaan stok hanya dilakukan dengan memeriksa persediaan barang secara manual. Sehingga hal ini kurang efisien jika konsumen membutuhkan barang dalam jumlah banyak dan stok tidak cukup untuk memenuhi kebutuhan konsumen.

Untuk mengatasi permasalahan tersebut diterapkan suatu metode data mining dengan cara menganalisa data penjualan untuk menghasilkan informasi yang dapat dijadikan sebagai perencanaan dan pengendalian persediaan barang. Adapun pengolahan datanya dapat dilakukan melalui proses clustering data dengan menerapkan metode k-means clustering. Metode k-means clustering bertujuan mengelompokkan data yang mempunyai karakteristik yang sama ke dalam satu cluster yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam cluster lain. Dengan dilakukan pengelompokan data tersebut dapat diketahui barang apa saja yang sudah terjual dan barang apa saja yang masih ada dalam stok.

Berdasarkan permasalahan di atas peneliti melakukan penelitian dengan judul “Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering”. Implementasi metode clustering tersebut lebih efektif dan efisien sehingga dapat tersedianya barang dengan jenis dan jumlah yang sesuai kebutuhan konsumen.

## 2. METODOLOGI PENELITIAN

Metode penelitian ini menggunakan Algoritma K-Means Clustering

### 2.1 Clustering

Clustering adalah metode pengelompokan data yang sering digunakan sebagai salah satu metode data mining atau penggalian data. Clustering adalah proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Oleh karena itu, metode clustering ini sangat berguna untuk menemukan kelompok yang tidak dikenal dalam data.

### 2.2 Algoritma K-Means

K-Means merupakan teknik clustering yang diperoleh dari sebuah dataset dengan cara menghitung jarak dari setiap titik ke pusat cluster secara iterative. Algoritma K-Means memiliki beberapa aturan yaitu:

1. Total cluster yang dibutuhkan
2. Jenis atribut adalah numeric. Proses untuk clustering menggunakan algoritma k-means adalah sebagai berikut :
  1. Tentukan nilai k dalam total cluster yang ingin dibentuk.
  2. Pilih nilai centroid cluster (centroid) berawal dari k.
  3. Hitung jarak setiap data input dan setiap centroid menggunakan rumus jarak Euclidean untuk menemukan jarak terdekat antara setiap titik data dan centroid.
  4. Mengklasifikasikan/mengelompokkan setiap item data berdasarkan jarak (jarak minimum) ke centroid.
  5. Perbarui mediannya. Nilai pusat baru.
  6. Mengulangi langkah ke 3 sampai 5 anggota setiap cluster tetap tidak berubah. Berikut ini merupakan rumus terhadap perhitungan algoritma K-Means Clustering :

$$de = \sqrt{(xi - si)^2 + (yi - ti)^2}$$

### 2.3 Metode Pengumpulan Data

Pengumpulan data dilakukan pada sebuah minimarket di daerah jambi pada bulan Desembertahun 2021. Barang dalam minimarket dikelompokkan kedalam 20. Data tersebut diolah untuk menciptakan pengetahuan yang dapat digunakan sebagai strategi untuk mengembangkan strategi pengelolaan persediaan barang. Data awal yang diperoleh dari minimarket disimpan sebagai datasemua penjualan obat selama 3 bulan terakhir.

Langkah-langkah prosedur menggunakan K-Means memiliki metode pengelompokanberikut:

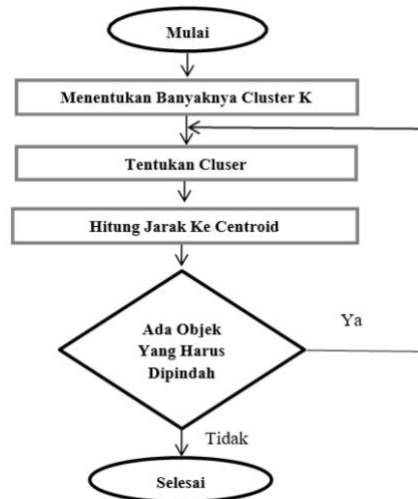
1. Tentukan nilai k untuk jumlah cluster yang ingin dibentuk.
2. Pilih pusat cluster awal (centroid) nilai k.
3. Hitung jarak dari setiap data input ke setiap centroid menggunakan rumus jarak Euclidean dantemukan jarak terdekat dari setiap data ke centroid.
4. Mengklasifikasikan/mengelompokkan setiap item data berdasarkan kedekatannya (jarakminimum) ke centroid.
5. Perbarui nilai pusat. Centroid baru diambil dari mean cluster tersebut menggunakan rumus.

Mengulangi langkah 3 sampai 5 anggota setiap cluster tetap tidak berubah. Dalam penelitian ini, mempergunakan K-Means Clustering untuk pengujian. Hal ini didasarkan pada penghitunganalgoritma K-Means Clustering dalam aplikasi yang dirancang untuk membuat keputusan yang digunakan untuk mengidentifikasi ketersediaan barang.

## 3 . HASIL PEMBAHASAN

### 3.1 Implementasi Algoritma K-Means

Ketika data siap di proses, algoritma K-Means memproses data pada gambar dibawah ini



Gambar 1. Tahapan Algoritma K-Means

Data penjualan barang ditampilkan dalam tabel dibawah ini:

Tabel 1. Data Penjualan barang

Data ke	Nama Obat	Satuan	Bulan Oktober	Bulan November	Bulan Desember
1	beras	kg	2	4	6
2	Gula	kg	5	1	2
3	minyak	kg	10	16	21
4	Shampo	botol	19	20	11
5	Tisu	box	31	17	15
6	garam	kg	49	60	91
7	deterjen	box	26	10	12
8	kecap	botol	15	81	16
9	saus	botol	9	10	11
10	sirup	botol	13	20	14
11	saus tiram	botol	16	17	20
12	air mineral	botol	60	36	51
13	wipol	botol	7	91	11
14	vanish	botol	20	21	16
15	so klin	sachet	11	86	4
16	rinso	sachet	30	49	100
17	daia	sachet	7	16	20
18	telur	Box	22	23	31
19	super pel	Botol	67	11	21
20	sabun mandi	Botol	79	49	100

### 3.2 Contoh perhitungan manual K-Means Clustering

Langkah-langkah dalam algoritma K-Means dijabarkan sebagai berikut:

1. mulai
2. menentukan jarak cluster k

setelah menentukan jumlah cluster pada penelitian ini, bagi cluster menjadi 2 cluster secara acak.

3. Studi penentuan Cluster, menurut Joanna Ardhiyanti M.N dan Yupie Kusumawati, penentuan focus awal. K-Means adalah acak. Oleh karena itu, dalam penelitian ini, pusat awal secara acak mengumpulkan dua data, Kelompok C1 dan C2, dan mengambil masing-

masing atribut. Sirup 13,20,14 dan gula 5, 1, 2. Menghitung jarak dari data ke pusat langkah selanjutnya adalah menghitung jarak data ke titik pusat menggunakan jarakEuclidean atau jarak terdekat untuk mendapatkan hasil data ke cluster C1 atau cluster C2.

Data ke 1: {2,4,6}

$$d_{euclidean}(C1) = \sqrt{(2 - 13)^2 + (4 - 20)^2 + (6 - 14)^2} = 21$$

$$d_{euclidean}(C2) = \sqrt{(2 - 5)^2 + (4 - 1)^2 + (6 - 2)^2} = 5,83095$$

Data ke 2: {3,2,2}

$$d_{euclidean}(C1) = \sqrt{(5 - 13)^2 + (1 - 20)^2 + (2 - 14)^2} = 23,85372$$

$$d_{euclidean}(C2) = \sqrt{(5 - 5)^2 + (1 - 1)^2 + (2 - 2)^2} = 0$$

Data ke 3: {10,16,21}

$$d_{euclidean}(C1) = \sqrt{(10 - 13)^2 + (16 - 20)^2 + (21 - 14)^2} = 8,60233$$

$$d_{euclidean}(C2) = \sqrt{(10 - 5)^2 + (16 - 1)^2 + (21 - 2)^2} = 24,71841$$

(Sampai data selanjutnya)

4. Apabila terdapat objek atau data yang berpindah maka ulangi perulangan hingga iterasi sampai tidak ada lagi yang berpindah.
5. Selesai.

Setelah dihitung sebagai iterasi pertama, dengan titik pusat diatas maka dihasilkan seperti tabel 2 dibawah ini:

Tabel 2. Menentukan pusat awal cluster

Bulan Oktober	Bulan November	Bulan Desember
13	20	14
5	1	2

Tabel 3. Menentukan jarak cluster

Dat a ke-	NamaObat	Satua n	Bulan Oktober	Bulan Novemb er	Bulan Desember	C1	C2	Jarak Terpendek
1.	beras	kg	2	4	6	21	5,83095189	5,83095189
2.	gula	kg	5	1	2	23,8537208	5	5
3.	minyak	kg	10	16	21	8,60232526	0	0
4.	shampo	Botol	19	20	11	6,70820393	24,7184141	8,60232526
5.	tisu	box	31	17	15	7	25,2586618	6,70820393
6.	garam	kg	49	60	91	18,2756668	33,1811320	18,2756668
7.	deterjen	box	26	10	12	8	05	8
8.	kecap	Botol	15	81	16	93,9414711	115,490259	93,9414711
9.	saus	botol	9	10	11	4	5	4
10.	sirup	Botol	13	20	14	16,5227116	24,9399278	16,5227116
11.	Saus tiram	botol	16	17	20	4	3	4
12.	Air mineral	botol	60	36	51	61,0655385	81,8290901	61,0655385
13.	wipol	botol	7	91	11	6	8	6
14.	vanish	botol	20	21	16	9	6	9
15.	So klin	sachet	11	86	4	11,1803398	13,3416640	11,1803398
16.	rinso	sachet	30	49	100	0	23,8537208	0
17.	daia	sachet	7	16	20	8	8	8
18.	telur	Box	22	23	31	7,34846922	26,4764045	7,34846922
19.	Super pel	Botol	67	11	21	8	9	8
20.	Sabun mandi	Botol	79	49	100	2	31	2
						71,3161971	90,4709898	71,3161973
						2	3	2
						7,34846922	28,6530975	7,34846922
						8	6	8
						66,7832314	85,2349693	66,7832314
						3	5	3
						92,3363417	111,950882	92,3363417
						1	1	1
						9,38083152	23,5159520	9,38083152
						3	3	3
						19,4679223	40,1746188	19,4679223
						3	5	3
						55,1905769	65,6124988	55,1905769
						1	1	1
						112,218536	131,848397	112,21853
						8	8	68

Tabel 4. Pengelompokan data

Data	C1	C2
1	1	0
2	1	0
3	1	0
4	1	0
5	1	0
6	1	0
7	1	0
8	0	1
9	1	0
10	1	0
11	1	0
12	1	0
13	0	1
14	1	0
15	0	0
16	1	1
17	1	0
18	1	0
19	1	0
20	1	0
n=	17	3

Pada perhitungan, dihasilkan dari 20 data terdapat 17 anggota dengan kategori C1 dan 3 anggota dengan kategori C2.

## 4. KESIMPULAN

Clustering (pengelompokan data) mempertimbangkan sebuah pendekatan penting untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. Clustering membagi kumpulan data ke dalam beberapa kelompok dimana kesamaan dalam sebuah kelompok adalah lebih besar daripada diantara kelompok-kelompok (Rui Xu dan Donald.2009). Gagasan mengenai pengelompokan data atau clustering, memiliki sifat yang sederhana dan dekat dengan cara berpikir manusia, kapanpun kepada kita dipresentasikan jumlah data besar ini ke dalam sejumlah kecil kelompok-kelompok atau kategori-kategori untuk memfasilitasi analisisnya lebih lanjut. Algoritma-algoritma clustering digunakan secara ekstensif tidak hanya untuk mengorganisasikan dan mengkategorikan data, akan tetapi juga sangat bermanfaat untuk kompresi data dan konstruksi model. Melalui pencarian kesamaan dalam data, seseorang dapat mempresentasikan data yang sama dengan lebih sedikit simbol. Pada hasil pengujian dengan 20 data, cluster optimal menyumbangkan 17 data untuk cluster C1 dan 3 data untuk cluster C2.

## UCAPAN TERIMAKASIH

Terimakasih peneliti sampaikan untuk Ibu Errisya Rasywir, S.Kom, MT Selaku dosenpengampu mata kuliah data mining dan selaku pembimbing atas selesainya penelitian ini. Terimakasih untuk pihak-pihak yang mendukung terlaksananya penelitian ini.

## REFERENCES

- [1] <https://finata.id/minimarket-dan-toko-kelontong-di-masyarakat/>
- [2] Nur, F., Zarlis, M., & Nasution, B. B. (2017). PENERAPAN ALGORITMA K-MEANS PADA SISWABARU SEKOLAHMENENGAH KEJURUAN UNTUK CLUSTERING JURUSAN. InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan), 1(2), 100-105.
- [3] RW Sembiring, JM Zain, A Embong "Clustering high dimensional data using subspace and projectedclustering algorithms" arXiv preprint arXiv:1009.0384