
PENERAPAN DATA MINING UNTUK MENGANALISIS PENGELOMPOKAN PEMBELIAN KONSUMEN MENGGUNAKAN ALGORITMA *K-MEANS CLUSTERING*

Cristi Girsang *, Marsono, **, Zaimah Panjaitan, **

* Sistem Informasi, STMIK Triguna Dharma

** Program Studi Dosen Pembimbing, STMIK Triguna Dharma

Article Info

Article history:

Received 12th, 201x

Revised 20th, 201x

Accepted 26th, 201x

Keyword:

First keyword

Second keyword

Third keyword

Fourth keyword

Fifth keyword

ABSTRACT

Clinic Distro merupakan perusahaan yang bergerak dibidang penjualan kebutuhan jasmani yang ada di Tanjung Morawa. Sampai saat ini Clinic Distro belum mempunyai sistem yang berbasis teknologi informasi yang digunakan sebagai sarana operasional perusahaan. Dengan banyaknya pelanggan dan data transaksi yang manual, perusahaan kesulitan dalam mendefinisikan karakteristik pelanggannya. Sehingga dibuatlah program Customer Relationship Management (CRM) dengan analisis segmentasi pelanggan berdasarkan RFM (Recency, Frequency, Monetary) menggunakan metode clustering dengan algoritma K-means. Metode penelitian yang digunakan pada penelitian ini tahapannya dimulai dari pengumpulan data. studi pendahuluan, pembuatan program, dan yang terakhir adalah pengujian cluster. Segmentasi pelanggan dengan metode clustering menghasilkan 3 cluster.

Copyright © 2020 STMIK Triguna Dharma.

All rights reserved.

Corresponding Author: *First Author

Nama :Cristi Girsang

Program Studi

STMIK Triguna Dharma

Email:cristigirsang13@gmail.com

1. PENDAHULUAN

Masyarakat perkotaan kini dimanjakan oleh kehadiran berbagai pusat perbelanjaan. Bahkan lokasinya kadang disatu kawasan. Kondisi ini tentu menguntungkan masyarakat sebagai konsumen karena masyarakat tinggal memilih gerai mana yang akan dimasukinya.

Dengan adanya persaingan ini diharapkan para pelaku usaha berkometisi secara kompetitif. Sehingga iklim usaha yang dihasilkan dapat berdampak positif kepada semua elemen yang terlibat di dalamnya. Ritel merupakan mata rantai terakhir proses distribusi antara barang dan jasa. Melalui ritel suatu produk akan langsung bertemu dengan penggunaannya. Sektor usaha ritel memiliki peranan yang cukup berpengaruh dalam perkembangan ekonomi di Indonesia. [1]

Hal ini terjadi karena perkembangan toko ritel masih menunjukkan tren positif. Toko ritel yang menjual berbagai pakaian menjadi primadona di kalangan para pengusaha. Sehingga para pengusaha menggelontorkan banyak modal untuk toko ritel mereka. Modal tersebut digunakan untuk menambah stok barang yang lebih kekinian dan modern. Di era modern sekarang, produsen dituntut untuk memahami pola perilaku konsumsi dari konsumen. Dengan diketahuinya pola konsumsi dari konsumen, para pelaku usaha dapat memprediksi dan menganalisa transaksi perbelanjaan. Sehingga para pelaku usaha lebih muda dalam melakukan aktivitas jual beli. Data mining sudah menjadi suatu piranti yang sangat populer dalam menangani banyak isi bisnis yang kompleks[2]

Hal ini dikarenakan kemampuan dari data mining untuk menyelesaikan permasalahan aktivitas usaha. Data mining memiliki sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan. Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses *Knowledge Discovery in Database* (KDD) yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya [3].

Tujuan dari penelitian ini adalah membuat suatu program atau aplikasi yang dapat mengelompokkan suatu barang yang paling diminati oleh konsumen berdasarkan data transaksi penjualan menggunakan metode *K-Means Clustering*. Maka dari itu, penelitian ini akan diangkat menjadi skripsi yang berjudul **“PENERAPAN DATA MINING UNTUK MENGANALISIS PENGELOMPOKAN PEMBELIAN KONSUMEN MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING.”**

2. METODE PENELITIAN

Penelitian yang baik sangat berpengaruh pada ketersediaan data yang diperlukan untuk pengambilan keputusan. Data yang tidak tepat berpengaruh terhadap keabsahan kesimpulan yang diambil. Untuk mendapatkan data akurat diperlukan cara penelitian yang tepat.

2.1 PENGUMPULAN DATA

1. Observasi

Observasi dilakukan bersama dengan kegiatan pelaksanaan sekaligus juga sebagai alat untuk pengumpulan data. Metode ini sesuai untuk merekam aktivitas yang bersifat proses. Kegiatannya seperti mengamati, merekam, dan mendokumentasikan setiap indikator.

2. Wawancara

Kegiatan wawancara dilakukan untuk mendapatkan informasi yang mendalam tentang pandangan, wawasan dari pihak-pihak yang memiliki wewenang agar memperoleh data yang dibutuhkan untuk mendukung penelitian ini.

Tabel 3.1 Data Penjualan *Clinic Distro* periode 2020 - 2021

Data Penjualan Per 2 Bulan Di Clinic Distro Periode 2020-2021					
No	Kode Barang	Nama Barang	Agustus-September	Oktober-November	Desember-Januari
1	A01	Tali Pinggang	15	13	5

2.2. Studi Kepustakaan (Study of Literature)

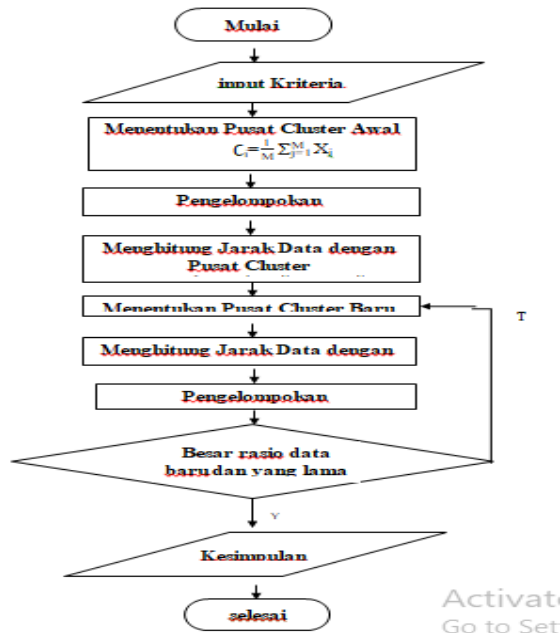
Dalam penelitian ini banyak menggunakan jurnal-jurnal baik jurnal nasional maupun buku sebagai sumber referensi. Dari komposisi yang ada jumlah literatur yang digunakan sebanyak 22 dengan rincian: 21 jurnal nasional, dan 1 buku nasional.

2.2. Algoritma Sistem

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan data mining dalam pengelompokan barang yang paling diminati pelanggan di *Clinic distro* berdasarkan kriteria dengan metode *K-Means*.

2.3 Flowchart Dari Metode Penyelesaian

Berikut ini adalah flowchart dari proses enkripsi dan dekripsi dari algoritma *K-Means Clustering* yaitu sebagai berikut



Gambar 3. 1 Flowchart algoritma K-Means Clustering

2.4 Dekripsi Data Penelitian

Berikut ini adalah data nasabah yang di dapat dari koperasi sinode parbu, yang akan diamankan. Dalam pengujiannya, sebagai contoh data yang digunakan sebagai sampel dalam penelitian ini yaitu sebagai berikut:

Tabel 3.2 Data Absensi Karyawan

NIK	Nama Karyawan	Tanggal	Jam Masuk	Jam Keluar	Keterangan
1166754267788	Prayitno	13/12/2020	9:30:00 AM	4:00:00 PM	-

2.4.1 Penyelesaian data dengan algoritma K-Means

Penetapan jumlah *cluster* (K) yaitu 3 *cluster*. Setelah menetapkan jumlah *cluster*, tentukan titik pusat awal *cluster* (*centroid*). Berikut ini titik *centroid* yang telah dipilih :

Tabel 3. 2 Titik Pusat Cluster awal

No	Kode Barang	Nama Barang	Data Penjualan Per 2 Bulan		
			Agustus-September	Oktober-November	Desember-Januari
1	A24	Kaos Ozon	3	114	80
2	A56	Sendal Niceman	3	12	18
3	A11	Sendal Thoste	2	2	1

1. Iterasi 1

a. Perhitungan jarak dengan pusat cluster dengan rumus

$$d(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_2 - x_{1j}|^2}$$

Tabel 3.3 Kelompok Cluster 1 Iterasi 1

KD	Agus-Sept	Okt-Nov	Des-Jan
A24	3	114	80
C1	3	114	80

Tabel 3. 1 Kelompok Cluster 2 Iterasi 1

KD	Agus-Sept	Okt-Nov	Des-Jan
A2	3	10	23
A3	7	12	42
A4	30	34	23
A6	13	17	25
A7	1	12	9
A10	3	12	12

Dari tabel diatas diperoleh *Centroid* baru yaitu :
 M1 (3,114,80), M2 (6.404,19.769,18.058) M3 (4,4.7,5.15)

Tabel 3. 7 Pusat *Centroid Cluster* Baru Iterasi 1

<i>Centroid 1</i>	3	114	80
<i>Centroid 2</i>	6.404	19.769	18.058
<i>Centroid 3</i>	4	4.7	5.15

2. Iterasi 2

Setelah menyelesaikan iterasi 1 maka dilanjutkan ke iterasi 2. Untuk lebih jelasnya dapat dilihat pada langkah-langkah berikut:

a. Perhitungan jarak dengan pusat *cluster* dengan rumus :

$$d(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_2 - x_{ij}|}$$

Tabel 3. 8 Jarak Data terhadap *cluster* pada iterasi kedua

KD	C1	C2	C3	Jarak Terdekat	JCT	JCT²
A1	132.853	21.506	11.005	C3	11.005	121.113
A2	118.596	11.465	18.647	C2	11.465	131.450
A3	108.922	25.178	37.686	C2	25.178	633.950

Tabel 3. 9 Kelompok *cluster* 1 Iterasi 2

KD	Agus-Sept	Okt-Nov	Des-Jan
A24	3	114	80
C1	3	114	80

Tabel 3. 10 Kelompok *cluster* 2 Iterasi 2

KD	Agus-Sept	Okt-Nov	Des-Jan
A2	3	10	23
A3	7	12	42
A4	30	34	23
A6	13	17	25

Tabel 3. 10 Kelompok *cluster* 2 Iterasi 2 (Lanjutan)

KD	Agus-Sept	Okt-Nov	Des-Jan
A14	4	1	45
A15	10	23	2
A23	1	35	18
A25	2	24	23

Tabel 3. 11 Kelompok *cluster* 3 Iterasi 2

KD	Agus-Sept	Okt-Nov	Des-Jan
A1	15	5	5
A5	2	4	10
A7	1	12	9

3. Iterasi 3

Setelah menyelesaikan iterasi 2 maka dilanjutkan ke iterasi 3. Untuk lebih jelasnya dapat dilihat pada langkah-langkah berikut:

- a. Perhitungan jarak dengan pusat *cluster* dengan rumus :

$$d(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_2 - x_{ij}|}$$

Tabel 3. 13 Jarak data terhadap *cluster* pada Iterasi ke 3

KD	C1	C2	C3	Jarak Terdekat	JCT	JCT²
A1	132.853	26.140	11.199	C3	11.199	125.416
A2	118.596	14.439	15.745	C2	14.439	208.471
A3	108.922	23.701	34.876	C2	23.701	561.748
A4	101.872	25.129	40.043	C2	25.129	631.471

4. Iterasi ke 4

Setelah menyelesaikan iterasi 3 maka dilanjutkan ke iterasi 4. Untuk lebih jelasnya dapat dilihat pada langkah-langkah berikut :

- a. Perhitungan jarak dengan pusat *cluster* dengan rumus :

$$d(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|}$$

Tabel 3. 20 Kelompok *Cluster* 1 Iterasi 4

KD	Agus-Sept	Okt-Nov	Des-Jan
A24	3	114	80
C1	3	114	80

Tabel 3. 2 Kelompok *Cluster* 2 Iterasi 4

KD	Agus-Sept	Okt-Nov	Des-Jan
A3	7	12	42
A4	30	34	23
A6	13	17	25
A14	4	1	45

Tabel 3. 3 Kelompok *Cluster* 2 Iterasi 4 (Lanjutan)

KD	Agus-Sept	Okt-Nov	Des-Jan
A62	2	23	20
A71	15	18	17
A72	10	23	17
C2	6.724	26.379	23.034

Tabel 3. 22 Kelompok *Cluster* 3 Iterasi 4

KD	Agus-Sept	Okt-Nov	Des-Jan
A1	15	5	5
A2	3	10	23
A5	2	4	10
A7	1	12	9
A8	4	4	7
A9	2	1	6

5. Iterasi 5

Setelah menyelesaikan iterasi 3 maka dilanjutkan ke iterasi 4. Untuk lebih jelasnya dapat dilihat pada langkah-langkah berikut:

a Perhitungan jarak dengan pusat *cluster* dengan rumus :

$$d(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|}$$

Tabel 3. 25 Kesimpulan Pengelompokan barang terlaris di Clinic Distro

KD	Nama Barang	Kelompok / Jarak Terdekat
A24	Kaos Ozon	C1
A3	Sendal RZZ	C2
A4	Topi	C2
A6	Boxer	C2

Keterangan :

Pengelompokan barang berdasarkan jumlah pembelian per 2 bulan di *Clinic Distro*. Untuk pusat *cluster* pertama dinyatakan ada 1 Merk Barang sedangkan untuk pusat *cluster* kedua dinyatakan ada 29 merk barang dan pada pusat *cluster* ketigadinyat akan ada 43 merk barang. Sesuai dengan *cluster* awal yang dipilih berdasarkan nilai rata-rata tertinggi, sedang dan yang terendah dari data maka berikut adalah susunan barang paling diminati:

C1 = *Cluster* Tertinggi dengan tingkat peminat paling banyak.

C2 = *Cluster* Menengah dengan tingkat peminat sedang.

C3 = *Cluster* Terendah dengan tingkat peminat rendah.

3. ANALISA DAN HASIL

3.1 Pengujian Sistem

Implementasi *data mining* untuk mengelompokkan pembelian di clinic distro pada periode agustus-september menggunakan metode *K-Means Clustering* membutuhkan 2 perangkat yaitu: Perangkat Lunak (*Software*) dan Perangkat Keras (*Hardware*).

1. Perangkat Lunak (*Software*)

Perangkat lunak yang dapat digunakan untuk sistem ini antara lain:

- a. Sistem operasi *Windows*
- b. *Microsoft Visual Basic Net 2012*
- c. *Microsoft Office Acces 2012*

2. Perangkat Keras (*Hardware*)

Perangkat keras yang dapat digunakan untuk sistem ini antara lain:

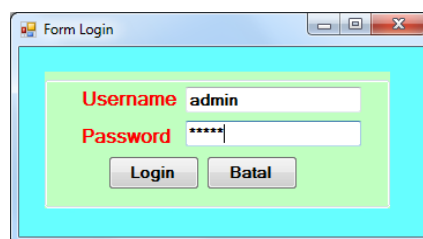
- a. *Prosessor Minimal Intel Dua Core Processor*
- b. *Hardisk Minimal 500 Gb*
- c. *Ram Minimal 2 Gb*
- d. *Monitor*

3.2 Implementasi Sistem

Implementasi merupakan langkah yang digunakan untuk mengoperasikan sistem yang dibangun. Dalam bab ini akan dijelaskan bagaimana menjalankan sistem tersebut. Dibawah ini merupakan tampilan dari aplikasi *data mining* untuk mengelompokkan barang terlaris berdasarkan tingkat pembelian periode Agustus–Januari di Clinic Distro Tanjung Morawa menggunakan metode *K-Means Clustering*.

1. Form Login

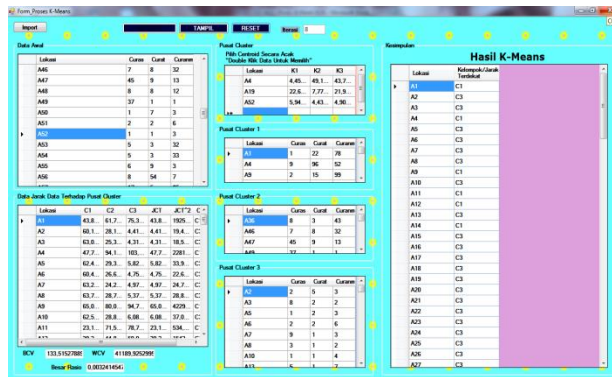
Form Login merupakan *form* yang digunakan sebagai media untuk membatasi hak akses. Cara menjalankannya adalah dengan mengisi *username* dan *password* yang benar kemudian klik tombol *login* untuk masuk kedalam sistem, klik tombol batal untuk mengosongkan *field*.



Gambar 5.1 Form Login

2. Form Proses

Form proses berisi tentang data perhitungan jarak data dengan pusat cluster serta pengelompokan setiap iterasi. cara menjalankannya adalah dengan meng-import data pembelian yang telah disediakan dalam bentuk file excel dengan cara menekan tombol import, kemudian pilih pusat cluster awal dengan cara double data yang di-import dimana data pusat cluster yang dipilih hanya 3 cluster. Setelah menentukan pusat cluster awal, klik tombol proses untuk melakukan perhitungan. Hasil perhitungan dapat dilihat dalam bentuk laporan dengan cara menekan tombol tampil. Klik tombol reset untuk mengosongkan semua field.



Gambar 5.2 Rancangan Form Proses

Gambar 5.3 Laporan

3. Form Laporan adalah form yang berisi tentang hasil akhir pengelompokan data pembelian di Clinic Distro pada periode Agustus- Januari.

CLINIC DISTRO		
Jl.Irian, Bandar Labuhan Tj. Morawa		
Laporan Pengelompokan Clinic Distro		
Lokasi	JT	Nama barang
A24	C1	kaos ozon
A3	C2	sendal rz2
A4	C2	topi
A6	C2	boxer
A14	C2	kaos linis biasa
A22	C2	tas samping
A23	C2	sepatu dokumento
A25	C2	kemeja panjang DRU
A26	C2	Sendal interco
A27	C2	kaos cloude
A28	C2	kaos de on
A29	C2	celana dalam
A30	C2	celana panjang becau:

Gambar 5.4 Laporan

4. KESIMPULAN

Dari hasil pembahasan bab I sampai bab V mengenai data mining untuk pembelian di Clinic Distro pada Periode Agustus-September dengan menggunakan algoritma *K-Means Clustering* dapat diambil beberapa kesimpulan sebagai berikut:

1. Sistem ini dibangun berdasarkan metode *K-Means Clustering* dan telah dilakukan perhitungan melalui data pembelian periode agustus-september. Sehingga diketahui barang terlaris pada periode tersebut .
2. Berdasarkan hasil analisa, pembelian terlaris di Clinic Distro pada periode Agustus-September ialah Produk Kaos Ozon.
3. Sistem yang digunakan dalam perhitungan *K-Means Clustering* ini ialah berbasis *desktop programming*.

UCAPAN TERIMA KASIH

Terimakasih buat keluarga dan para sahabat yang selalu menemani dan memberikan suport kepada saya selama dalam menyelesaikan artikel ilmiah ini terkhusus buat kedua orang tua saya yang mendoakan agar saya dapat berhasil dan teman-teman yang tiada henti-hentinya memberikan masukan-masukan dan saran dalam penyusunan karya ilmiah ini.

REFERENSI

- [1] Zubi Mahrofi, "No Title," *Antaraneews.com*, 2019. <https://www.antaraneews.com/berita/1159256/mendag-industri-ritel-miliki-peran-penting-dalam-pertumbuhan-ekonomi>.
- [2] Verasius Albert, "No Title," *Binus University*, 2019. <https://binus.ac.id/malang/2019/01/penerapan-penerapan-data-mining-seri-data-mining-for-business-intelligence-4/>.
- [3] P. Bidang, K. Sains, Y. Mardi, J. Gajah, M. No, and S. Barat, "Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika."
- [4] C. Informasi, "Informasi," 2014.
- [5] S. Elemen and R. Pranala, "Sistem," pp. 1–3.
- [6] hendini dan andi, "Pemodelan UML Sistem Informasi Monitoring Penjualan Stok Barang," *J. Khatulistiwa Inform.*, vol. 5, pp. 107–116, 2016.
- [7] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2017, doi: 10.30743/infotekjar.v2i1.164.
- [8] M. A. W. K. MURTI, "Penerapan Metode K-Means Clustering Untuk Mengelompokan Potensi Produksi Buah – Buah Di Provinsi Daerah Istimewa Yogyakarta," *Skripsi*, 2017.

- [9] S. Handoko, E. Tri, and E. Handayani, "IMPLEMENTASI DATA MINING UNTUK MENENTUKAN TINGKAT PENJUALAN PAKET DATA TELKOMSEL MENGGUNAKAN METODE K-MEANS CLUSTERING," vol. 25, no. 1.
- [10] F. L. Sibuea, A. Sapta, S. Informasi, and S. Royal, "PEMETAAN SISWA BERPRESTASI MENGGUNAKAN METODE K-MEANS CLUSTERING," vol. IV, no. 1, 2017.
- [11] J. Tech and A. K-means, "Analisis Data Mining Pada Strategi Penjualan Produk PT Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering," vol. 2, no. 1, pp. 32–41, 2019.
- [12] Y. Syahra, "Penerapan Data Mining Dalam Pengelompokan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma

BIBLIOGRAFI PENULIS

	<p>Nama :Cristi Girsang Nirm : 2017020877 Tempat/tgl lahir : Medan , 10 juli 1996 Jenis Kelamin : Perempuan Alamat : jl.sultan serdang pasar 5 gg. Datuk Agama : Kristen Protestan Email : cristigirsang13@gmail.com No hp : 085377964103</p>
	<p>Nama : Marsono, S.Kom., M.Kom NIDN : 0102057501 Jenis Kelamin : Laki-Laki Email : marsonotgdsi3@gmail.com</p>
	<p>Nama : Zaimah Panjaitan, S.Kom., M.Kom NIDN : 0120098903 Tempat/tgl lahir : - Jenis Kelamin : Perempuan Email : zaimahpanjaitan@gmail.com No hp : 08137034099 Pendidikan : - S1- STMIK Triguna Dharma -S2 – Universitas Putra Indonesia Yptk Padang Bidang Keahlian : Keamanan Komputer, Artificial Intelligence,dll</p>