

Penerapan Data Mining Untuk Mengelompokkan Data Penjualan XL Home Menggunakan Metode K-Means

Nanda Guslan Bakri¹, Azanuddin², Fifi Sonata³

¹Sistem Informasi, STMIK Triguna Dharma

²Teknik Komputer, Politeknik Negeri Medan

³Manajemen Informatika, STMIK Triguna Dharma

Email: ¹nandaguslan22@gmail.com, ²azdin.bpc@gmail.com, ³fifinsonata2012@gmail.com

Email Penulis Korespondensi: nandaguslan22@gmail.com

Abstrak

Pada PT. GSR memiliki karyawan sebanyak 18 Orang sales yang memasarkan produk XL Home di kota Medan. Dalam waktu 1 bulan, target penjualan diharapkan sebanyak 200 pelanggan baru didapatkan untuk berlangganan produk XL Home. Tetapi, dengan luasnya cakupan area penjualan tidak sesuai dengan jumlah sales yang dimiliki. Permasalahan tersebut perusahaan membutuhkan bidang keilmuan data mining dengan menggunakan metode K-Means Clustering. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama. Hasil penelitian ini dapat pengelompokan data penjualan berdasarkan target penjualan dan jumlah penjualan di kota medan ini menggunakan data mining dengan metode clustering diharapkan dapat membantu untuk PT. GSR dalam melakukan promosi penjualan ke wilayah yang memiliki potensi besar dengan jumlah sales yang telah dimiliki agar menjadi lebih efisien dan mampu mencapai target penjualan yang diberikan.

Kata Kunci: Data Mining, K-Means, Penjualan, XL Home, karyawan.

Abstract

PT. GSR has 18 sales employees who market XL Home products in the city of Medan. Within 1 month, the sales target is expected to be 200 new customers to subscribe to XL Home products. However, the wide coverage of the sales area does not correspond to the number of sales owned. The problem is that the company needs a scientific field of data mining using the K-Means Clustering method. By using this method, the data that has been obtained can be grouped into several clusters based on the similarity of the data, so that data that has the same characteristics is grouped in one cluster and those that have different characteristics are grouped in another cluster that has the same characteristics. The results of this research can group sales data based on sales targets and the number of sales in the city of Medan using data mining with the clustering method, which is expected to be helpful for PT. GSR carries out sales promotions to areas that have great potential with the number of sales it already has so that it becomes more efficient and able to achieve the given sales targets.

Keywords: Data Mining, K-Means, Sales, XL Home, employees.

1. PENDAHULUAN

Setiap perusahaan mempunyai data yang tersimpan dalam basis datanya. Data transaksi tersebut semakin hari semakin banyak dan bertambah. Dengan bertambahnya jumlah data pada perusahaan tersebut, maka peran analisis untuk menganalisis data secara manual perlu digantikan dengan aplikasi yang berbasis komputer. Sehingga proses penganalisis dapat dilakukan secara tepat dan akurat.

Data mining adalah analisa terhadap suatu data untuk menentukan hubungan yang jelas serta menyimpulkan yang belum diketahui sebelumnya dengan cara terkini dan berguna bagi pemilik data tersebut. Teknik data mining dibagi menjadi beberapa kelompok sesuai tugas yang dilakukan yaitu deskripsi, estimasi, prediksi, klasifikasi, klastering dan asosiasi. Melihat dari permasalahan-permasalahan yang sudah dipaparkan sebelumnya, maka teknik data mining untuk menyelesaikan masalah yang ada dengan mengolah data historis transaksi penjualan yang sudah ada. Dalam permasalahan yang ada penulis menggunakan metode K-Means, dengan menklaster atau mengelompokkan Produk yang terjual.

Pada PT. GSR yang bergerak di bidang penjualan Produk XL Home kota Medan memiliki permasalahan dalam melakukan pengelompokan data- data pelanggan potensial di kota medan. Pelanggan yang tersebar di kota medan memiliki karakteristik yang berbeda-beda untuk setiap wilayahnya [1].

Pada PT. GSR memiliki karyawan sebanyak 18 Orang sales yang memasarkan produk XL Home di kota Medan. Dalam waktu 1 bulan, target penjualan diharapkan sebanyak 200 pelanggan baru didapatkan untuk berlangganan produk XL Home. Tetapi, dengan luasnya cakupan area penjualan tidak sesuai dengan jumlah sales yang dimiliki. Dikarenakan keterbatasan itu maka diperlukan strategi pengelompokan wilayah yang memiliki potensi penjualan.

Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). Dengan data mining, kita dapat melakukan pengklasifikasian, memprediksi, memperkirakan dan mendapatkan informasi lain yang bermanfaat dari kumpulan data dalam jumlah yang besar. Salah satu metode yang terdapat dalam *data mining* yang digunakan untuk

penelitian ini adalah metode *clustering* karena merupakan metode yang sudah cukup dikenal dan banyak dipakai dalam *data mining* [2].

Tujuan utama dari metode *clustering* ini adalah pengelompokan sejumlah data/obyek ke dalam *cluster* (*group*) sehingga dalam *cluster* akan berisi data yang sama dengan grupnya masing-masing. Dalam *clustering* metode ini berusaha untuk menempatkan obyek yang mirip (jaraknya dekat) dalam satu *cluster* dan membuat jarak antar *cluster* sejauh mungkin. Digunakannya metode *clustering* pada data penjualan mesin dan bahan-bahan digital printing untuk menjadikan data tersebut *cluster-cluster* [3]. Sehingga nantinya pengelompokan data penjualan berdasarkan target penjualan dan jumlah penjualan di kota medan ini menggunakan data mining dengan metode *clustering* diharapkan dapat membantu untuk PT. GSR dalam melakukan promosi penjualan ke wilayah yang memiliki potensi besar dengan jumlah sales yang telah dimiliki agar menjadi lebih efisien dan mampu mencapai target penjualan yang diberikan.

2. METODOLOGI PENELITIAN

2.1 Data mining

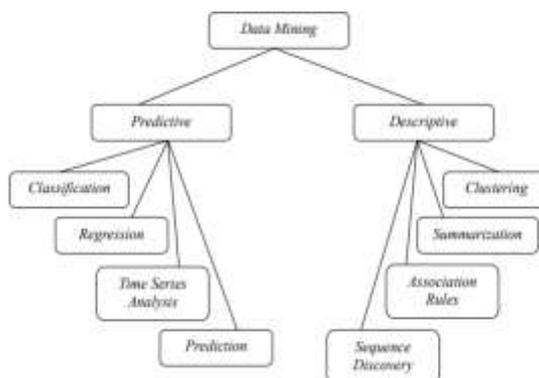
Data mining adalah sebuah proses menganalisa data untuk mengetahui suatu pola dari kelompok data yang tersembunyi [4]. Yang bertujuan untuk menemukan, menggali pengetahuan dari data atau informasi yang kita miliki. *Data mining* melibatkan pencarian pola yang diperlukan dalam database untuk membantu pembuat keputusan di masa yang akan datang.

Data Mining adalah kenyataan bahwa Data Mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari beberapa disiplin ilmu [5] dan Tujuan utama data mining adalah untuk menemukan, atau menambang pengetahuan dari data atau informasi yang kita miliki.

Data mining, sering juga disebut sebagai bagian dari *knowledge discovery in database* (KDD). KDD merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam *set data* berukuran besar. Pada tahapan data mining ini dilakukan proses ekstraksi data yang besar untuk menemukan pengetahuan yang belum tersampaikan sebelumnya, dengan menggunakan teknik klasifikasi [6].

2.2 Teknik Data mining

Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan [7]. Teknik data mining biasanya dibagi menjadi dua kategori yaitu prediksi dan deskripsi. Teknologi peramalan menggunakan data historis untuk menyimpulkan informasi tentang peristiwa di masa depan. Sedangkan teknik deskriptif bertujuan untuk menemukan pola dalam data yang memberikan beberapa informasi tentang hubungan interval yang tersembunyi [8].



Gambar 1 Teknik *Data mining*

2.3 Metode K-Means dengan teknik Clustering

Algoritma *K-means* merupakan salah satu algoritma dengan partitional, karena *K-Means* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya [9]. Algoritma *K-means* menggunakan proses secara berulang-ulang untuk mendapatkan basis data *cluster* [10]. Algoritma *K-Means* adalah metode pengelompokan berbasis jarak yang membagi data ke dalam sejumlah *cluster* [11], serta memiliki kemudahan dan kemampuan untuk menganalisa data besar dengan cepat dan dapat menemukan data *outlier*. Langkah-langkah *clustering* menggunakan metode *K-means* adalah sebagai berikut [12]:

1. Pilih jumlah *cluster* *k*.
2. Alokasikan data ke dalam *Cluster* secara random
3. Hitung *centroid* atau rata-rata dari data yang ada di masing-masing *cluster*

Alokasikan masing-masing data ke *centroid* atau rata-rata terdekat

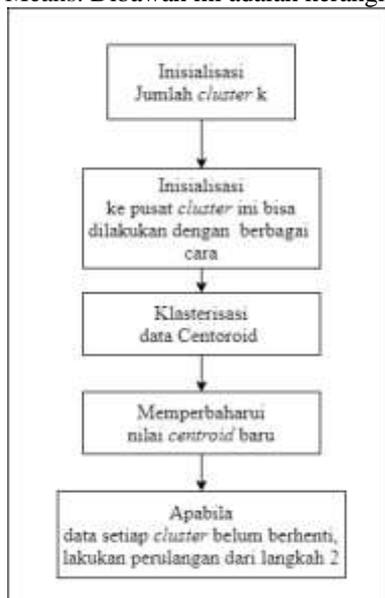
3. HASIL DAN PEMBAHASAN

3.1 Algoritma Sistem

Algoritma adalah sekumpulan aturan yang secara tepat menentukan urutan operasi. *Algoritma System* adalah suatu urutan ataupun tahapan-tahapan dalam proses pembuatan system dimana akan memberikkan keluaran yang dikehendaki berdasarkan masukan yang diberikan.

3.1.1 Kerangka Kerja Algoritma K-Means

Kerangka kerja algoritma merupakan tahapan operasi proses algoritma yang digunakan dalam menyelesaikan masalah dengan menggunakan algoritma K-Means. Dibawah ini adalah kerangka kerja dari algoritma *K-Means*.



Gambar 2. Kerangka Kerja Algoritma K-Means

Adapun tahapan-tahapan yang dilakukan dalam proses pembuaatan sistem adalah sebagai berikut.

1. Inisialisasi Jumlah *cluster* k.
2. Inisialisasi ke pusat *cluster* ini bisa dilakukan dengan berbagai cara. Cara yang paling sering dilakukan adalah dengan *random* atau acak. Pusat-pusat *cluster* diberi dengan nilai awal dengan angka-angka *random*.

$$d(x_i, \mu_j) = \sqrt{\sum (x_i, \mu_j)^2}$$

Dimana:

x_i = data kriteria

μ_j = *centroid* pada *cluster* ke- j_s

3. Klasterisasi data Centoroid

Berdasarkan kedekatannya dengan *centroid* atau mencari jarak terkecil.

4. Memperbaharui nilai *centroid* baru

Nilai *centroid* baru di peroleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus yaitu :

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j$$

Keterangan :

$\mu_j(t+1)$ = *centroid* baru pada *iterasi* $(t+1)$ N_{sj} = Data pada *cluster* S_j

5. Apabila data setiap *cluster* belum berhenti, lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah Menampilkan hasil klasterisasi penjualan data

3.2 Penerapan Metode *K-Means*

Berikut ini adalah deskripsi data yang diuji setelah dilakukan pengumpulan data melalui wawancara dengan sales gudang PT.GSR pada September 2020 -September 2021.

Tabel 1 Penjualan September 2020 -September 2021

No	Nama Sales	Target	Terjual
1	Adi Pranata Brahmana	330	315
2	Ana Linda Br Ginting	58	53
3	Anisa Br Sembiring	48	45
4	Amri Surbakti	202	202
5	Amika Putri Br Sitepu	172	166
6	Amira Kartika	90	81
7	Apriansyah Putra	102	92
8	Billy Sembiring	81	81
9	Boby Siregar	63	57
10	Bunga Rafelia Br Barus	95	81
11	Elisabet Silalahi	64	58
12	Evan Saputra	151	146
13	Handika P. Tarigan	170	162
14	Handika P. Tarigan	133	128
15	Joel Sion Sembiring	140	132
16	Josua Ginting	330	325
17	Jessica Wati	121	115
18	Kalia Pepayosa Br Ginting	40	36
19	Kristian Saragih	26	25
20	Mita Wulandari	170	156
21	Mailida Nehe	31	22
22	Masita Br Bangun	20	19
23	Maya sartika	151	140
24	Meylinda sari	30	27
25	Rendi pemalemsa sembiring	22	20
26	Reza diansyah	90	88
27	Rere nika br keliat	70	62
28	Rido manta kemit	25	20
29	Rifaldo baru	31	29
30	Revi chania sembiring	21	18
31	Wawan baslius keliat	25	21
32	Sakira venita	20	18
33	Dani marvelanda surbakti	52	50
34	Ananta jaya ginting	92	81
35	Angelita clodia	80	77
36	Anike br pandia	36	34
37	Anisa lila br ketaren	43	41
38	Arya ginting	80	73
39	Beny andika bukit	35	36
40	Boby persadanta baru	55	52
41	Cristian p. Sitepu	36	31
42	Cristy anastasya br pinem	30	25
43	Cindya br sembiring	20	19
44	Fransiscus da diva	35	34
45	Friska liana	460	402
46	Hosea oktavianus	65	61
47	Henri paldo	50	42
48	Heru sinulinnga	45	39
49	Iwanta sahputra bukit	62	55
50	Icha wandita br ginting	57	44
51	Jesica anjani	35	32
52	Liska sari br ginting	45	37
53	Nindy natasya br baru	173	160
54	Nelsi br purba	30	24

55	Revaldo tarigan	45	41
56	Risa ninta br sitepu	160	156
57	Rima br tarigan	31	27
58	Silvani br keliat	40	35
69	Oska gurusinga	33	28
60	Saskia laura	40	35
61	Paula ninta	43	41
62	Pitra maulana	50	47
63	Widya mayangsari	58	51
64	Wira kesuma	44	40
65	Nurul Silvani	40	35
66	Putri Wahyuni	34	27
67	Ragil Suhardina	56	48
68	Rama Eka Satria	40	39
69	Ramadani	35	33
70	Rizka Amelia Azis	22	20
71	Silvy Dwi Tamara	36	34
72	Siswanto	177	177
73	Sri Agustina Br Tarigan	152	140
74	Tasya Awalia Sitepu	153	150
75	Tri Wulandari	201	201
76	Yunita Sari Sembiring	35	32
77	Zahilul Muhibah	40	35
78	Afnida Aulia Husna Sigalingging	72	61

3.2.1 Menentukan Inisialisasi Jumlah Cluster Dan Titik Awal Cluster

Sebelum melakukan perhitungan terlebih dahulu harus menentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang ada akan dikelompokkan menjadi tiga *cluster*. Selanjutnya adalah menentukan titik pusat awal dari setiap *cluster*. Dalam penelitian ini titik pusat *cluster* yang dipilih dapat dilihat pada tabel 3.2 sebagai berikut:

Tabel 2. Contoh Data Titik Pusat *Cluster* pada iterasi-1

Titik Pusat Awal	No Data	Nama Sales	Pemesanan	Penjualan
<i>Cluster 1</i>	45	Friska liana	460	402
<i>Cluster 2</i>	34	Sakira venita	92	81
<i>Cluster 3</i>	32	Ananta jaya ginting	20	18

3.2.2 Inisialisasi Ke Pusat Cluster

Inisialisasi ke pusat *cluster* dengan menghitung jarak data ke *Centroid* menggunakan rumus *Euclidean*, data tersebut dari *Cluster* terdekatnya.

- a. Menghitung data pertama dengan titik *cluster 1*

$$C1 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C1 = \sqrt{(330-460)^2 + (315-402)^2}$$

$$C1 = 156,426$$

- b. Menghitung data pertama dengan titik *cluster 2*

$$C2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C2 = \sqrt{(330-92)^2 + (315-81)^2}$$

$$C2 = 333,766$$

- c. Menghitung data pertama dengan titik *cluster 3*

$$C3 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$C3 = \sqrt{(330-20)^2 + (315-18)^2}$$

$$C3 = 429,312$$

3.2.3 Klasterisasi penjualan Setiap Data Centoroid

Berikut tabel dibawah ini dapat dilihat untuk hasil lengkap dari perhitungan pada tahap iterasi pertama:

Tabel 3 Hasil Perhitungan Iterasi ke-1

No	Kode Sales	C1	C2	C3	Jarak Terdekat	Berada Di Cluster
1	A001	156,426	333,766	429,312	156,426	1
2	A002	532,358	44,045	51,662	44,045	2
3	A003	545,154	56,851	38,897	38,897	3
4	A004	326,441	163,527	258,805	163,527	2
5	A005	372,344	116,726	212,151	116,726	2
6	A006	489,838	2,000	94,175	2,000	2
7	A007	473,565	14,866	110,454	14,866	2
8	A008	496,671	11,000	87,693	11,000	2
9	A009	525,960	37,643	58,052	37,643	2
10	A010	486,072	3,000	97,949	3,000	2
11	A011	524,549	36,235	59,464	36,235	2
12	A012	401,269	87,784	183,153	87,784	2
13	A013	376,431	112,450	207,933	112,450	2
14	A014	426,620	62,370	157,699	62,370	2
15	A015	418,688	70,036	165,517	70,036	2
16	A016	151,093	340,852	436,290	151,093	1
17	A017	444,173	44,688	140,036	44,688	2
18	A018	557,096	68,768	26,907	26,907	3
19	A019	574,878	86,556	9,220	9,220	3
20	A020	380,284	108,208	203,823	108,208	2
21	A021	573,098	84,865	11,705	11,705	3
22	A022	583,343	95,016	1,000	1,000	3
23	A023	405,123	83,439	179,011	83,439	2
24	A024	570,548	82,219	13,454	13,454	3
25	A025	581,178	92,849	2,828	2,828	3
26	A026	485,279	7,280	98,995	7,280	2
27	A027	517,397	29,069	66,603	29,069	2
28	A028	578,921	90,609	5,385	5,385	3
29	A029	568,480	80,156	15,556	15,556	3
30	A030	583,247	94,921	1,000	1,000	3
31	A031	578,261	89,939	5,831	5,831	3
32	A032	584,000	95,671	0,000	0,000	3
33	A033	538,858	50,606	45,255	45,255	3
34	A034	488,329	0,000	95,671	0,000	2
35	A035	500,025	12,649	84,149	12,649	2
36	A036	561,427	73,110	22,627	22,627	3
37	A037	551,552	63,253	32,527	32,527	3
38	A038	502,634	14,422	81,394	14,422	2
39	A039	560,875	72,622	23,431	23,431	3
40	A040	535,280	47,011	48,795	47,011	2
41	A041	563,398	75,073	20,616	20,616	3
42	A042	571,864	83,546	12,207	12,207	3
43	A043	583,343	95,016	1,000	1,000	3
44	A044	562,182	73,878	21,932	21,932	3
45	A045	0,000	488,329	584,000	0,000	1
46	A046	521,829	33,601	62,241	33,601	2
47	A047	545,619	57,315	38,419	38,419	3
48	A048	551,357	63,032	32,650	32,650	3
49	A049	528,027	39,699	55,973	39,699	2
50	A050	539,048	50,931	45,222	45,222	3
51	A051	563,494	75,166	20,518	20,518	3
52	A052	552,675	64,382	31,401	31,401	3
53	A053	375,410	113,146	208,741	113,146	2

54	A054	572,524	84,220	11,662	11,662	3
55	A055	550,042	61,717	33,971	33,971	3
56	A056	387,964	101,237	196,581	101,237	2
57	A057	569,795	81,468	14,213	14,213	3
58	A058	557,754	69,426	26,249	26,249	3
59	A059	567,631	79,310	16,401	16,401	3
60	A060	557,754	69,426	26,249	26,249	3
61	A061	551,552	63,253	32,527	32,527	3
62	A062	542,333	54,037	41,725	41,725	3
63	A063	533,671	45,343	50,329	45,343	2
64	A064	551,453	63,127	32,558	32,558	3
65	A065	557,754	69,426	26,249	26,249	3
66	A066	567,539	79,246	16,643	16,643	3
67	A067	537,152	48,836	46,861	46,861	3
68	A068	555,130	66,843	29,000	29,000	3
69	A069	562,837	74,518	21,213	21,213	3
70	A070	581,178	92,849	2,828	2,828	3
71	A071	561,427	73,110	22,627	22,627	3
72	A072	361,544	128,222	223,450	128,222	2
73	A073	404,361	84,149	179,744	84,149	2
74	A074	397,181	92,098	187,385	92,098	2
75	A075	327,844	162,114	257,391	162,114	2
76	A076	563,494	75,166	20,518	20,518	3
77	A077	557,754	69,426	26,249	26,249	3
78	A078	516,551	28,284	67,476	28,284	2

Dari tabel 3.4 diatas dapat hasil iterasi ke-1 sebagai berikut:

- C1= {A001,A016 dan A045}
- C2={ A002, A004, A005, A006, A007, A008, A009, A010, A011, A012, A013, A014, A015, A017, A020, A023, A026, A027, A034, A035, A038, A040, A046, A049, A053, A056, A063, A072, A073, A074, A075, A078}
- C3={ A002, A004, A005, A006, A007, A008, A009, A010, A011, A012, A013, A014, A015, A017, A020, A023, A026, A027, A034, A035, A038, A040, A046, A049, A053, A056, A063, A072, A073, A074, A075, A078}

1. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 156,426 + 44,045 + 38,897 + 163,527 + \dots + 28,284$$

$$WCV = 3171,189$$

2. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} \text{a. } D(m1,m2) &= \sqrt{(m1-m2)^2} \\ &= \sqrt{(460-92)^2+(402-81)^2} \\ &= 488,329 \end{aligned}$$

$$\begin{aligned} \text{b. } D(m1,m3) &= \sqrt{(m1-m3)^2} \\ &= \sqrt{(460-20)^2+(402-18)^2} \\ &= 95,671 \end{aligned}$$

$$\begin{aligned} \text{c. } D(m2-m3) &= \sqrt{(m2-m3)^2} \\ &= \sqrt{(92-20)^2+(81-18)^2} \\ &= 584,000 \end{aligned}$$

$$\begin{aligned} \text{Nilai BCV} &= d(m1,m2) + d(m1,m3)+d(m2,m3) \\ &= 488,329 + 95,671 + 584,000 \\ &= 1168,000 \end{aligned}$$

3. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV.

$$\begin{aligned} BCV/WCV &= 1168,000 / 3171,189 \\ &= 0,368 \end{aligned}$$

3.2.4 Memperbarui Nilai *Centroid* Baru

Kemudian langkah berikutnya pada iterasi ke-2 adalah Menghitung kembali pusat-pusat baru dari *cluster* yang baru terbentuk dengan menghitung nilai rata-rata pada *cluster*.

1. Pada *cluster* 1 terdapat 3 data maka nilai rata-rata pada *cluster*-1 adalah

$$s = (330 + 330 + 460) / 3 = 373,333$$

$$t = (315 + 325 + 402) / 3 = 347,333$$

- Pada *cluster 2* terdapat 9 data maka nilai rata-rata pada *cluster-2* adalah
 $s = (58 + 202 + 172 + 90 + 102 + 81 + 63 + 95 + 64 + 151 + 170 + 133 + 140 + 121 + 170 + 151 + 90 + 70 + 92 + 80 + 80 + 55 + 65 + 62 + 173 + 160 + 58 + 177 + 152 + 153 + 201 + 72) / 32 = 115,718$
 $t = (53 + 202 + 166 + 81 + 92 + 81 + 57 + 81 + 58 + 146 + 162 + 128 + 132 + 115 + 156 + 140 + 88 + 62 + 81 + 77 + 73 + 52 + 61 + 55 + 160 + 156 + 51 + 177 + 140 + 150 + 201 + 61) / 32 = 109,218$
- Pada *cluster 3* terdapat 67 data maka nilai rata-rata pada *cluster-3* adalah
 $s = (48 + 40 + 26 + 31 + 20 + 30 + 22 + 25 + 31 + 21 + 25 + 20 + 52 + 36 + 43 + 35 + 36 + 30 + 20 + 35 + 50 + 45 + 57 + 35 + 45 + 30 + 45 + 31 + 40 + 33 + 40 + 43 + 50 + 44 + 40 + 34 + 56 + 40 + 35 + 22 + 36 + 35 + 40) / 43 = 36,093$
 $t = (45 + 36 + 25 + 22 + 19 + 27 + 20 + 20 + 29 + 18 + 21 + 18 + 50 + 34 + 41 + 36 + 31 + 25 + 19 + 34 + 42 + 39 + 44 + 32 + 37 + 24 + 41 + 27 + 35 + 28 + 35 + 41 + 47 + 40 + 35 + 27 + 48 + 39 + 33 + 20 + 34 + 32 + 35) / 43 = 32,209$

Tabel 4 Titik Pusat *Cluster* pada iterasi-2

TITIK PUSAT AWAL	PENJUALAN	PENJUALAN
<i>Cluster 1</i>	373,333	347,333
<i>Cluster 2</i>	175,333	170,000
<i>Cluster 3</i>	36,093	32,209

3.3 Interface (antarmuka)

Fungsi dari *interface* (antarmuka) ini adalah untuk memberikan *input* dan menampilkan *output* dari aplikasi. Pada aplikasi ini memiliki *interface* yang terdiri dari *Menu login*, *Data Sales*, *Data Centroid* dan *Menu Proses K-Means*.

3.3.1 Halaman Utama

Dalam halaman utama untuk menampilkan pada tampilan *Menu* pada awal sistem yaitu *Menu login* dan *Menu Utama*. Adapun *Menu* halaman utama sebagai berikut.

1. *Menu Login*

Menu Login digunakan untuk mengamankan sistem dari *user-user* yang tidak bertanggung jawab sebelum masuk ke *Menu Utama*. Berikut adalah tampilan *Menu Login* :



Gambar 2 *Menu Login*

2. *Menu Utama*

Menu Utama digunakan sebagai penghubung untuk *Menu Data Sales*, *data Centroid*, proses dan laporan. Berikut adalah tampilan *Menu Utama*.



Gambar 3 *Menu Utama*

3.3.2 Halaman Administrator

Administrator untuk menampilkan *Menu* pengolahan data pada penyimpanan data kedalam *database* yaitu *Menu* Data Sales, dan *Menu* Centroid. Adapun *Menu* halaman *administrator* utama sebagai berikut.

1. Menu Data Sales

Menu data Sales berfungsi untuk pengolahan dalam penginputan data, ubah data dan penghapusan data Sales. Adapun *Menu* data nasabah adalah sebagai berikut.



No	Nama Sales	Kode Sales	C1
1	Anisa Br Sembiring	1	48
2	Amri Surbakti	2	202
3	Amika Putri Br Sitepu	3	172
4	Amira Kartika	4	90
5	Apriansyah Putra	5	102
6	Billy Sembiring	6	81
7	Boby Siregar	7	63

Gambar 4 Menu Data Sales

2. Menu Data Centroid

Menu Data Centroid untuk pengolahan data Centroid pemilihan produk asuransi. Adapun *Menu* Data Centroid adalah sebagai berikut.



Centroid - 1	460	402
Centroid - 2	92	81
Centroid - 3	20	18

Gambar 5 Menu Data Centroid

3.3.3 Pengujian

Pada bagian ini diminta untuk melakukan pengujian dengan *sampling* data baru untuk dapat menguji keakuratan sistem yang dirancang dengan *tools-tools* yang sudah teruji dan terkalibrasi sebelumnya. Adapun hasil proses program dalam penjualan bahan pokok sebagai berikut.



The screenshot shows the K-MEANS application interface. At the top, there is a table with columns: No, Nama, Kode, C1, and C2. Below this table are buttons for 'PROSES' and 'CETAK LAPORAN', and a field for 'Iterasi' set to 'Iterasi-3'. Below the buttons is a section titled 'Cluster baru' containing a table with columns: No, Nama, Cluster-1, Cluster-2, and Cluster.

No	Nama	Kode	C1	C2
1	Anisa Br Sembiring	1	48	45
2	Amri Surbakti	2	202	202
3	Amika Putri Br Sitepu	3	172	166
4	Amira Kartika	4	90	81
5	Apriansyah Putra	5	102	92
6	Billy Sembiring	6	81	81
7	Boby Siregar	7	63	57
8	Bunga Rafelia Br Barus	8	95	81

No	Nama	Cluster-1	Cluster-2	Cluster
1	Anisa Br Sembiring	48	45	Kurang
2	Amri Surbakti	202	202	Cukup
3	Amika Putri Br Sitepu	172	166	Cukup
4	Amira Kartika	90	81	Cukup
5	Apriansyah Putra	102	92	Cukup
6	Billy Sembiring	81	81	Cukup
7	Boby Siregar	63	57	Kurang
8	Bunga Rafelia Br Barus	95	81	Cukup
9	Elisabet Silalahi	64	58	Kurang
10	Evan Saputra	151	146	Cukup

Gambar 6 Hasil Analisa K-Means



The screenshot shows a report window titled 'Laporan_Cluster'. The report header includes 'PT. GSR' and 'JL PAMAH / JL PUTRI HIJAU DELI TUA (DELI SERDANG), Medan, Indonesia 20147'. The main content is a table titled 'Laporan Hasil Pengelompokan' with columns: Kode Sales, Nama Sales, and Hasil.

Kode Sales	Nama Sales	Hasil
1	Anisa Br Sembiring	Kurang Laris
2	Amri Surbakti	Cukup Laris
3	Amika Putri Br Sitepu	Cukup Laris
4	Amira Kartika	Cukup Laris
5	Apriansyah Putra	Cukup Laris
6	Billy Sembiring	Cukup Laris
7	Boby Siregar	Kurang Laris
8	Bunga Rafelia Br Barus	Cukup Laris

Gambar 7 Laporan Hasil

4. KESIMPULAN

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang dibahas tentang mengelompokkan data perusahaan dengan menerapkan metode terhadap *k-means* sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan dalam menganalisa pengelompokkan segepentasi *sales* dilakukan dengan riset ke tempat perusahaan dan

wawancara pihak pengolahan data 78 data sales yang dikelompokkan dalam segementasi penjualan dan Dalam merancang ataupun membangun sistem menggunakan bahasa pemodelan *Unified Modeling Language* (UML) yang terdiri dari use case *diagram*, *class diagram*, *activity diagram*. Pembangunan sistem dengan bahasa pemograman *visual basic*.

Dalam akurasi metode yang digunakan sangat tepat berdasarkan jumlah data digunakan 78 dan menampilkan hasil pengelompokkan data penjualan dengna cepat dan akurat.

UCAPAN TERIMAKASIH

Terima Kasih diucapkan kepada kedua orang tua serta keluarga yang selalu memberi motivasi, Doa dan dukungan moral maupun materi, serta pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya jurnal ini bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal selanjutnya.

DAFTAR PUSTAKA

- [1] D. Sunia and P. Alam Jusia, "PENERAPAN DATA MINING UNTUK CLUSTERING DATA PENDUDUK MISKIN MENGGUNAKAN ALGORITMA K-MEANS".
- [2] S. Syahidatul Helma, R. R. Rustiyan, E. Normala, P. Studi Sistem Informasi Fakultas Sains dan Teknologi, U. Islam Negeri Sultan Syarif Kasim Riau, J. Soebrantas No and S. Baru, "Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means," 2019.
- [3] F. Yunita, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTRING PADA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS ISLAM INDRAGIRI)," 2018.
- [4] D. Nofriansyah, M. Yetri, K. Erwansyah, and _ S., "Penerapan Data Mining Dalam Menganalisa Data Penjualan Untuk Mendapatkan Pola Rekomendasi Penjualan Menggunakan Algoritma Apriori Pada K3 Mart," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 2, p. 176, 2019, doi: 10.53513/jis.v18i2.157.
- [5] S. M. Riniarti and S. Bahri, "Produktivitas Karet Pada Lahan Hkm Jaya Lestari Kabupaten Way Kanan Provinsi Lampung Rubber Productivity On Community Forest Jaya Lestari, Way Kanan District, Lampung Province," *Gorontalo Journal of Forestry Research*, vol. I, no. 1, pp. 37-43, 2018.
- [6] A. A. Fajrin and A. Maulana, "Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor," *Kumpulan jurnaL Ilmu Komputer (KLIK)*, vol. V, no. ISSN: 2406-7857, pp. 27-36, 2018.
- [7] Y. Kristyawan and L. P. Sumirat, "ANALISIS TERHADAP FAKTOR-FAKTOR YANG MEMPENGARUHI MAHASISWA GAGAL STUDI MENGGUNAKAN TEKNIK KLASIFIKASI," 2019.
- [8] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. V, no. 2407-0491, pp. 213-219, 2018.
- [9] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," *Jurnal Teknik Industri*, vol. X, no. 2622-5131, pp. 275-282, 2020.
- [10] D. H. Makalalag, Y. D. Y. Rindengan and R. Sengkey, "Sistem Informasi Geografis Monitoring Dan Evaluasi Proyek Dinas Pekerjaan Umum Kota Kotamobagu," *E-journal Teknik Informatika*, vol. IX, no. 1, pp. 1-9, 2016.
- [11] J. Hutagalung and F. Sonata, "Penerapan Metode K-Means Untuk Menganalisis Minat Nasabah," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 1187, 2021, doi: 10.30865/mib.v5i3.3113.
- [12] M. T. Prihandoyo, "Unified Modeling Language (UML) Model Untuk Pengembangan Sistem Informasi Akademik Berbasis Web," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. III, no. 2477-5126, pp. 126-129, 2018 .