



## **IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN ARKETIPE PEMBELIAN SUKU CADANG DAN ASESORIS KOMPUTER(STUDI KASUS DI TOKO LAKSAMANA KOMPUTER DUMAI)**

**Fauzansyah**

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Dumai

Jl. Utama Karya Bukit Batrem, Dumai, Kode Pos : 28811

e-mail: fauzansyah3@gmail.com

### **ABSTRAK**

*Penelitian ini dilakukan untuk mempelajari Data Mining K-Means Clustering dalam pencarian knowledge (pengetahuan). Tujuan dari penelitian ini kemungkinan dapat membantu pihak Laksamana Komputer Dumai untuk menentukan asesoris dan spare part computer mana yang pembeli paling banyak dan diletakkan pada posisi paling depan berdasarkan permintaan konsumen di Laksamana Komputer Dumai. Untuk itu dalam metode K-Means Clustering dimungkinkan adanya solusi dan analisa terhadap pengolahan data dan parameter-parameter yang menjadi acuan untuk mengambil keputusan. Di dalam metode ini terdapat langkah-langkah penyelesaian masalah. Adapun tools bantu untuk mengimplementasikan metode tersebut adalah Rapid miner 7.3. Rapid miner 7.3 akan mengolah data secara tersusun atas operator-operator yang nestable yang langsung didapatkan hasil secara akurat selanjutnya pada tahapan terakhir akan didapatkan knowledge baru.*

**Kata Kunci :** *Data Mining, Aloritma K-Means, Knowledge, Rapid Miner 7.3.*

## **1. PENDAHULUAN**

### **1.1. Latar Belakang**

Dengan kemajuan perekonomian yang sangat pesat pada saat sekarang ini membuat orang – orang berlomba-lomba untuk meningkatkan penjualan produknya. Untuk peningkatan produk yang di jualnya banyak terdapat kesalahan dalam hal menyiapkan laporan yang disampaikan kepada atasannya terhadap penjualan yang terus terjadi, dan tidak sedikit karyawan yang mendapat teguran dari atasannya. Untuk mencapai hal itu, ada tiga kebutuhan bisnis yang dapat dilakukan, yaitu penambahan jenis

produk maupun meningkatkan kapasitas produk, pengurangan operasional perusahaan, dan peningkatan efektifitas pemasaran serta keuntungan. Sementara para pelaku bisnis memiliki kebutuhan-kebutuhan untuk memanfaatkan “*gudang data*“ yang sudah dimilikinya, para peneliti melihat peluang itu untuk melahirkan sebuah teknologi baru yang menjawab kebutuhan ini, yaitu *data mining* (Nurjoko dan Darmawan A., 2015).

*Data mining* mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Namun banyak nya data yang ada , tidak selalu



dibarengi dengan pengetahuan yang dapat dihasilkan oleh data yang banyak tersebut, sehingga pada akhirnya data-data tersebut hanya menjadi suatu yang kurang berguna. *Data mining* merupakan proses pencaharian pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, seperti data *were house* dan tempat penyimpanan data lainnya (Ramadhani A., Farmadi A. dan Budiman I., 2014).

Objek-objek/data yang dikelompokkan ke dalam suatu grup memiliki ciri-ciri yang sama berdasarkan kriteria tertentu, salah satu aktifitas analisis data adalah klasifikasi atau pengelompokan data kedalam beberapa kategori atau *cluster*. Untuk melakukan pengelompokan terhadap data pengelompokan suku cadang dan asesoris komputer menggunakan teknik *clustering*. Metode yang digunakan *K-Means Clustering*. Dengan menggunakan metode ini data-data yang didapatkan dapat dikelompokkan ke dalam beberapa *cluster*, berdasarkan minat beli konsumen, sehingga data-data yang dimiliki karakteristik yang sama dikelompokkan dalam satu *cluster* yang lain (Mulyati S., 2015).

Dalam hal ini Toko Laksamana Komputer Dumai menyediakan berbagai macam *SparePark* dan Asesoris Komputer untuk keperluan konsumen seperti *keyboard*, *mouse*, *motherboard*, memori atau RAM dan lain sebagainya. Semakin banyak nya konsumen yang selalu memesan barang-barang yang disediakan membuat karyawannya tidak sempat melakukan pencatatan terhadap persediaan barang tersebut khususnya di Toko Laksamana Komputer Dumai.

Dari permasalahan tersebut di atas, maka penulis tertarik melakukan penelitian ini dengan judul: “Implementasi Data Mining Menggunakan Metode Algoritma K-Means Clustering Untuk Menentukan Arketipe Pembelian Suku Cadang Dan Asesoris Komputer (Studi Kasus Di Toko Laksamana Komputer Dumai)”.

**1.2. Tujuan Penelitian**

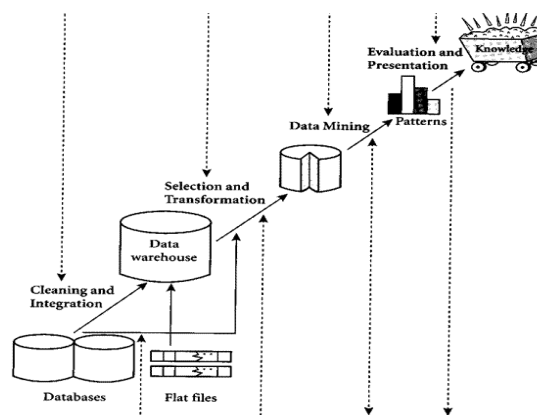
Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Menganalisa setiap karakteristik produk data yang telah diperoleh dari Toko Laksamana Komputer Dumai.
2. Mengelompokkan data ke dalam *cluster* yang berbeda secara iteratif .
3. Untuk menentukan arketipe pembelian asesoris dan *sparepart* komputer dengan menggunakan *algoritma k-means clustering*.

**1.3. Tinjauan Pustaka**

*Knowledge Discovery in Database (KDD)* adalah keseluruhan proses *non-trivial* untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola ditemukan bersifat sah, baru dapat bermanfaat dan dimengerti. Pada proses *KDD (Knowledge Discovery and Data Mining)* sering kali merupakan aplikasi iteratif yang berulang dari metodologi data mining tertentu. Tujuan utama dari *KDD* adalah untuk membuat prediksi dan deskripsi. Prediksi menggunakan beberapa variabel atau *field-field* basis data untuk memperediksi nilai-nilai variabel masa mendatang yang diperlukan, yang belum diketahui saat ini (Ndaumanu R. I., Kusriani, & Arief M. R., 2014).

Data mining adalah bagian dari *Knowledge Discovery in Database (KDD)* yaitu sebuah proses mencari pengetahuan yang bermanfaat dari data, proses *KDD* secara garis besar dapat dijelaskan sebagai berikut (Sunjana., 2010).



Gambar 1. Tahapan-tahapan *Data Mining*



Clustering adalah membagi data kedalam grup-grup yang mempunyai objek yang karakteristiknya sama tertentu (Aras Z., and Sarjono., 2016). Proses mengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/*cluster*. Tujuannya menemukan *cluster* yang berkualitas dalam waktu yang layak. *Clustering* dalam *data mining* berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data. Kesamaan objek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi.

*K-Means Clustering* adalah suatu metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi (Nasari F., dan Darma S., 2015).

## 2. METODE PENELITIAN

### 2.1. Kerangka Penelitian

Adapun urutan langkah kerja dari penelitian sebagai berikut :

#### 1. Mengidentifikasi Masalah

Masalah yang diidentifikasi dalam penelitian ini adalah besaran jumlah penjualan produk yang laris terjual dan data master yang diambil dari Toko Laksamana Komputer Dumai masih acak, serta menggali informasi terpendam dari banyaknya barang yang berkemungkinan memiliki nilai lebih. Oleh karena itu penulis ingin menganalisa serta mengelompokkan informasi – informasi tersebut dengan menggunakan *metode k-means clustering*.

#### 2. Menganalisa Masalah

Bagaimana data penjualan yang sering dikeluhkan dapat diolah menggunakan metode *metode k-means clustering*. Sehingga dapat diperoleh pengetahuan baru untuk mendukung

keputusan pada Toko Laksamana Komputer Dumai.

#### 3. Menentukan Tujuan

Berdasarkan uraian di atas, tujuan yang akan dicapai dari penelitian ini adalah:

- Untuk memberikan masukan kepada Toko Laksamana Komputer Dumai berupa ide - ide berdasarkan pengetahuan yang telah didapat.
- Untuk mendapatkan *knowledge* (pengetahuan) tentang penjualan produk dan asesoris komputer.
- Untuk membuat data yang cepat, akurat dan memberikan informasi berdasarkan hasil yang didapat dari penelitian ini.

#### 4. Memahami Literatur

Literatur-literatur yang dipakai sebagai bahan referensi dalam penelitian ini adalah dari jurnal-jurnal ilmiah nasional dan yang membahas tentang *data mining* khususnya tentang metode *Algoritma K-Means Clustering*. Literatur - literatur ini akan menjadi pedoman untuk melakukan penelitian agar mempermudah dalam proses penelitian.

#### 5. Mengumpulkan Data

Metode pengumpulan data adalah cara-cara dan prosedur yang dilakukan dalam penelitian untuk memperoleh data yang dibutuhkan. Beberapa metode pengumpulan data yang dilakukan dalam penelitian ini adalah sebagai berikut :

- Melakukan *observasi* dan kajian tentang data transaksi penjualan Toko Laksamana Komputer Dumai. Kajian lapangan ini untuk mengetahui secara langsung permasalahan yang ada, sehingga diharapkan penerapan disain dan konsep analisa data Perusahaan menggunakan algoritma *K-Means* dapat dilakukan dengan tepat.
- Library Research* (tinjauan kepustakaan) dilakukan untuk mengumpulkan informasi tentang literatur dan pedoman dalam penentuan hasil analisa data Perusahaan menggunakan algoritma *K-Means*.



c. *Laboratory Research* (Penelitian Laboratorium) adalah untuk pengujian data penentuan hasil analisa data *Perusahaan* yang telah didapatkan dengan menggunakan algoritma *K-Means*.

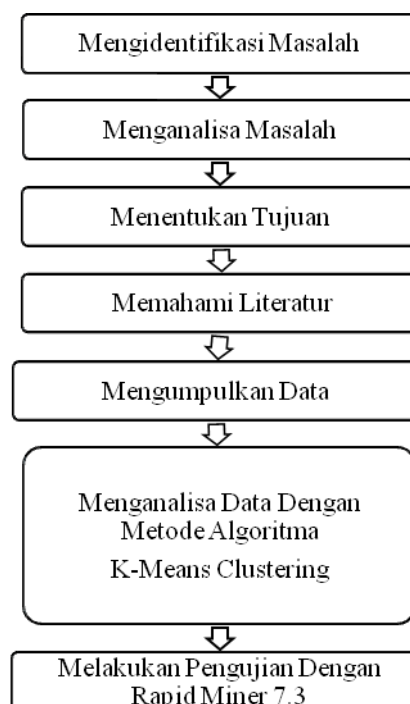
6. Menganalisis Data Dengan Metode Algoritma K-Means Clustering

Data yang telah dikumpulkan akan dianalisa. Dimana data – data tersebut dikelompokkan menjadi, masalah pribadi, masalah sosial, masalah perusahaan, dan permasalahan karir. Pada tahap ini juga dilakukan tahapan *cleaning*, yaitu tahapan dimana data-data yang telah dikelompokkan untuk kemudian dilakukan pembersihan data. Adapun cara pembersihan data tersebut adalah dengan cara melengkapi data, menghapus data duplikat, Dan menghilangkan *noise*, setelah dilakukan *cleaning*, Maka dilanjutkan dengan transformasi data-data yang telah dibersihkan pada tahap sebelumnya. Adapun cara transformasi data ini adalah dengan cara memformat data, Sehingga siap di cluster atau dikelompokkan.

7. Melakukan Pengujian Dengan Aplikasi *Rapid Miner 7.3*.

Evaluasi sistem yang telah di uji perlu dilakukan, dengan tujuan untuk mengetahui apakah sistem yang diuji tersebut sudah sesuai dengan yang diharapkan. Berikut ini adalah mekanisme pengujian yang dilakukan :

1. Pengujian manual dengan menggunakan metode *Algoritma K-Means Clustering*. Di mana dalam mencari nilai dari variabel input dilakukan proses *Cluster* dengan mencari jumlah anggota tiap *cluster*.
2. Pengujian dengan menggunakan *software Rapid Miner 7.3*. untuk mendapatkan hasil yang lebih optimal dan kesalahan dalam proses pengambilan keputusan dapat diminimalkan.
3. Hasil (a) dibandingkan dengan hasil (b) untuk validasi.



Gambar 2. Kerangka Kerja Penelitian

**3. HASIL DAN PEMBAHASAN**

Ada pun langkah-langkah yang dilakukan pada tahapan *K-Means Clustering* adalah sebagai berikut (M.Agustin F.E., et al., 2015) :

1. Tentukan k sebagai jumlah *cluster* yang ingin dibentuk. Tetapkan pusat *cluster*.
2. Hitung jarak setiap data ke pusat *cluster* menggunakan persamaan Euclidean.

$$D(i_k) = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \dots \dots \dots (1)$$

3. Kelompokkan data ke dalam *cluster* yang dengan jarak yang paling pendek menggunakan persamaan.

$$\text{Min} \sum_{k=1}^k dik = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \dots \dots \dots (2)$$

4. Hitung pusat *cluster* yang baru menggunakan persamaan.

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \dots \dots \dots (3)$$

Dimana:

$X_{ij}$  e *cluster* ke - k



- p = banyaknya anggota *cluster* ke k.
- Ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

Tabel 1. Sampel Data Rekapitulasi Penjualan Toko Laksamana Komputer

No	Nama Barang	Banyak Terjual	Harga/Unit
1	Handset gaming	76	412
2	Mouse wereless	41	192
3	Cd blank/pac	59	269
4	Kabel vga 15 m	60	299
5	Kabel hdmi 10m	61	291
6	Converter hdmi to vga	75	474
7	Vga splitter 2 port	82	397
8	Vga splitter 4 port	172	974
9	Cd blank/pac	123	716
10	Tinta canon 1 liter black	71	339
11	Tinta canon 1 liter colour	58	300
12	Kabel usb 1 pack	69	500

**3.1 Analisa pengelompokan data**

Dalam menentukan pengelompokan data atau banyak nya jumlah *cluster* yang akan dibuat, dapat dilakukan beberapa pertimbangan teoritis dan konseptial yang mungkin diusulkan untuk menentukan beberapa banyak *cluster* yang harus dibentuk. Pada penelitian ini penentuan jumlah kelompok atau *cluster* ditentukan berdasarkan pertimbangan pada pengelompokan data penjualan asesoris dan *sparepart* komputer. Dimana variabel yang digunakan adalah sebanyak 2 variabel, yaitu harga barang dan total penjualan untuk satu tahun.

**3.2 Analisa dan perancangan proses algoritma k-means clustering**

Pada penjelasan sebelumnya terdapat sampel data sebanyak 12 produk yang dijual dan digunakan untuk melakukan percobaan perhitungan manual, Ada pun proses sebagai berikut :

**1. Iterasi 1**

- Pusat awal *cluster* atau *centroid* diambil berdasarkan tingkat penjualan tinggi, sedang dan rendah dari sampel data rekapitulasi penjualan toko laksamana komputer pada item banyak produk yang terjual di asumsikan :

Tabel 2 Pusat Awal Cluster

Di ambil data ke-8 sebagai pusat cluster ke-1	172	974
Di ambil data ke-7 sebagai pusat cluster ke-2	82	387
Di ambil data ke-2 sebagai pusat cluster ke-3	41	192

**2. Perhitungan jarak pusat cluster**

Untuk mengukur jarak antara data dengan pusat cluster digunakan rumus *Euclidian distance*, kemudian akan didapatkan matrik jarak sebagai berikut:

Rumus *Euclidian distance* :  $D(i,k)$   
 $= \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$

Dimana :

$C_{ij}$  = Pusat cluster

$C_{kj}$  = Data

Dari 12 data yang dijadikan sampel telah dipilih pusat awal *cluster* yaitu **C1(172,974)**, **C2(82,397)** dan **C3(41,192)**. Lalu dilakukan perhitungan jarak dari sisa sampel data dengan pusat *cluster* yang dimisalkan dengan **M(a,b)**, dimana **a** merupakan total penjualan, dan **b** harga barang yang diperkecil menjadi angka puluhan juta agar cara penghitungan lebih mudah.

- M1 = (76,412)
- M2 = (41,192)
- M3 = (59,269)
- M4 = (60,299)
- M5 = (61,291)
- M6 = (75,474)
- M7 = (82,397)
- M8 = (172,974)
- M9 = (123,716)
- M10 = (71,339)
- M11 = (58, 300)
- M12 = (69,500)



Hitung *Euclidian distance* dari semua data ketiap titik pusat pertama :

$$\begin{aligned}
 D1.1 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(76 - 172)^2 + (412 - 974)^2} = 570,14 \\
 D1.2 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(41 - 172)^2 + (192 - 974)^2} = 792,8 \\
 D1.3 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(59 - 172)^2 + (269 - 974)^2} = 713,9 \\
 D1.4 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(60 - 172)^2 + (299 - 974)^2} = 684,2 \\
 D1.5 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(61 - 172)^2 + (291 - 974)^2} = 691,9 \\
 D1.6 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(75 - 172)^2 + (474 - 974)^2} = 509,3 \\
 D1.7 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(82 - 172)^2 + (397 - 974)^2} = 583,9 \\
 D1.8 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(172 - 172)^2 + (974 - 974)^2} = 0 \\
 D1.9 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(123 - 172)^2 + (716 - 974)^2} = 262,6 \\
 D1.10 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(71 - 172)^2 + (339 - 974)^2} = 642,9 \\
 D1.11 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(58 - 172)^2 + (300 - 974)^2} = 683,5 \\
 D1.12 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(69 - 172)^2 + (500 - 974)^2} = 485,06
 \end{aligned}$$

Dengan cara yang sama hitung jarak tiap titik ketitik pusat ke-2 dan kita akan mendapatkan :

$$\begin{aligned}
 D2.1 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(76 - 82)^2 + (412 - 397)^2} = 16,1 \\
 D2.2 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(41 - 82)^2 + (192 - 397)^2} = 209,05 \\
 D2.3 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(59 - 82)^2 + (269 - 397)^2} = 130,05 \\
 D2.4 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(60 - 82)^2 + (299 - 397)^2} = 100,4 \\
 D2.5 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(61 - 82)^2 + (291 - 397)^2} = 108,06 \\
 D2.6 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(75 - 82)^2 + (474 - 397)^2} = 77,3 \\
 D2.7 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(82 - 82)^2 + (397 - 397)^2} = 0 \\
 D2.8 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(72 - 82)^2 + (974 - 397)^2} = 583,9 \\
 D2.9 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(123 - 82)^2 + (716 - 397)^2} = 321,6 \\
 D2.10 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(71 - 82)^2 + (339 - 397)^2} = 59,03 \\
 D2.11 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(58 - 82)^2 + (300 - 397)^2} = 99,9 \\
 D2.12 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(69 - 82)^2 + (500 - 397)^2} = 103,8 \\
 \text{Dengan cara yang sama hitung jarak tiap titik ketitik pusat ke-3 dan kita akan mendapatkan :} \\
 D3.1 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(76 - 41)^2 + (412 - 192)^2} = 222,7 \\
 D3.2 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(41 - 41)^2 + (192 - 192)^2} = 0 \\
 D3.3 &= \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \\
 &= \sqrt{(59 - 41)^2 + (269 - 192)^2} = 79,07
 \end{aligned}$$



$$D3.4 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(60 - 41)^2 + (299 - 192)^2} = 108,6$$

$$D3.5 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(61 - 41)^2 + (291 - 192)^2} = 101$$

$$D3.6 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(75 - 41)^2 + (474 - 192)^2} = 284,04$$

$$D3.7 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(82 - 41)^2 + (397 - 192)^2} = 209,06$$

$$D3.8 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(172 - 41)^2 + (974 - 192)^2} = 792,8$$

$$D3.9 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(123 - 41)^2 + (716 - 192)^2} = 530,3$$

$$D3.10 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(71 - 41)^2 + (339 - 192)^2} = 150,03$$

$$D3.11 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(58 - 41)^2 + (300 - 192)^2} = 109,3$$

$$D3.12 = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} = \sqrt{(69 - 41)^2 + (500 - 192)^2} = 399,2$$

3. Perhitungan pengelompokan data ke dalam cluster dengan jarak yang paling pendek.

Untuk mengukur jarak terpendek data dengan pusat cluster digunakan rumus jarak terpendek sebagai berikut:

$$\text{Min} \sum_{k=1}^k \text{dik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

atau = MIN (Cij : Ckj : Clj )

Untuk melakukan perhitungan jarak terpendek dengan pusat cluster data diambil dari hasil perhitungan D1,D1,D3 adalah sebagai berikut :

$$J1.1 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(570,1:16,1:222,7) = 16,1$$

$$J1.2 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(792,8:209,05:0) = 0$$

$$J1.3 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(713,9:130,05:79,07) = 79,07$$

$$J1.4 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(684,2:100,4:108,6) = 100,4$$

$$J1.5 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(691,9:108,06:101) = 101$$

$$J1.6 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(509,32:77,3:284,04) = 77,3$$

$$J1.7 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(583,9:0:209,06) = 0$$

$$J1.8 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(0:583,9:792,8) = 0$$

$$J1.9 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(262,6:321,6:530,3) = 262,6$$

$$J1.10 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(642,9:59,03:150,03) = 59,03$$

$$J1.11 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(683,5:99,9:109,3) = 99,9$$

$$J1.12 = \text{MIN} (Cij : Ckj : Clj) = \text{MIN}(485,06:103,8:309,2) = 103,8$$

4. Alokasikan data-data ke centroid yang paling dekat

setelah diketahui hasil setiap jarak *euclidean distance* dan jarak terdekat maka didapatkan hasil iterasi 1 dan pengelompokan data sebagai berikut:

Tabel 3. Hasil Perhitungan Iterasi 1

No	Nama Barang	C1	C2	C3	Jarak Terpendek
1	Handset gaming	570.1403	16.15549	222.7667	16.15549442
2	Mouse wereless	792.8966	209.0598	0	0
3	Cd blank/pac	713.9986	130.05	79.07591	79.07591289
4	Kabel vga 15 m	684.2288	100.439	108.6738	100.4390362
5	Kabel hdmi 10m	691.961	108.0602	101	101
6	Converter hdmi to vga	509.3221	77.31753	284.0423	77.31752712
7	Vga splitter 2 port	583.9769	0	209.0598	0
8	Vga splitter 4 port	0	583.9769	792.8966	0
9	Cd blank/pac	262.6119	321.624	530.3772	262.6118809
10	Tinta canon 1 liter black	642.9821	59.03389	150.03	59.03388857



11	Tinta canon 1 liter colour	683.573	99.92497	109.3298	99.92497185
12	Kabel usb 1 pack	485.0619	103.8171	309.2701	103.8171469

Tabel 4 Hasil Pengelompokan Iterasi 1

No.	C1	C2	C3
1		1	
2			1
3			1
4		1	
5			1
6		1	
7		1	
8	1		
9	1		
10		1	
11		1	
12		1	

Hasil nilai 1 pada tabel diatas sebagai petunjuk pengelompokan data jarak dan data jarak *auclidean distance* yang mempunyai nilai sama.

**2. Iterasi ke-2**

1. Hitung pusat *centroid* baru tentukan posisi *centroid* baru (Ck) dengan cara menghitung nilai rata-rata yang ada pada *centroid* yang sama dengan rumus sebagai berikut :

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p}$$

Dimana nk adalah jumlah dokumen dalam cluster k dan d1 adalah dokumen dalam *cluster* k. Sehingga didapatkan titik pusat atau *centroid* yang baru yaitu:

Tabel 5 Perhitungan Pusat Baru

No	Nama barang	banyak terjual	harga/unit
1	Handset gaming	76	412
2	Mouse wereless	41	192
3	Cd blank/pac	59	269
4	Kabel vga 15 m	60	299
5	Kabel hdmi 10m	61	291
6	Converter hdmi to vga	75	474
7	Vga splitter 2 port	82	397

Tabel 6 Perhitungan Pusat Baru (Lanjutan)

8	Vga splitter 4 port	172	974
9	Cd blank/pac	123	716
10	Tinta canon 1 liter black	71	339
11	Tinta canon 1 liter colour	58	300
12	Kabel usb 1 pack	69	500

**3.3 Analisa dan hasil**

Berdasarkan analisa dan hasil pengujian implementasi data mining menggunakan algoritma k-means clustering untuk menentukan arketipe pembelian suku cadang dan asesoris di toko laksana komputer yang dilakukan telah dapat Memberikan gambaran dalam pengambilan keputusan serta memberikan informasi dari data penjualan yang diolah dari sampel yang telah diuji.

Tabel 7 Data Sampel Penjualan Asesoris Dan *Sparepart* Komputer

No	Nama Barang	Banyak Terjual	Harga/unit
1	Handset gaming	76	412
2	Mouse wereless	41	192
3	Cd blank/pac	59	269
4	Kabel vga 15 m	60	299
5	Kabel hdmi 10m	61	291
6	Converter hdmi to vga	75	474
7	Vga splitter 2 port	82	397





8	Vga splitter 4 port	172	974
9	Cd blank/pac	123	716
10	Tinta canon 1 liter black	71	339
11	Tinta canon 1 liter colour	58	300
12	Kabel usb 1 pack	69	500

Tabel 8 Hasil Sampel Pengelompokan Pengujian Menggunakan K-Means Clustering

No.	C1	C2	C3
1		1	
2			1
3			1
4			1
5			1
6		1	
7		1	
8	1		
9	1		
10		1	
11			1
12		1	

Setelah diketahui hasil sampel pengelompokan tabel data diatas bahwa angka 1 menandakan posisi pengelompokan data berdasarkan cluster yang sudah diproses. Kemudian angka tersebut diganti berdasarkan produk atau asesoris agar mempermudah dalam pengelompokan yang akan ditunjukkan pada tabel dibawah ini.

Tabel 9 Hasil Sampel Pengelompokan Asesoris Dan Produk Di Toko Laksamana Komputer

No.	Pengelompokan 1	Pengelompokan 2	Pengelompokan 3
1		Handset gaming	
2			Mouse wereless
3			Cd blank/pac
4			Kabel vga 15 m
5			Kabel hdmi 10m
6		Converter hdmi to vga	
7		Vga splitter 2 port	
8	Vga splitter 4		

	port		
9	Cd blank/pac		
10		Tinta canon 1 liter black	
11			Tinta canon 1 liter colour
12		Kabel usb 1 pack	

Berdasarkan tabel diatas proses dan pengujian dengan beberapa iterasi yang telah dilakukan, maka bahwa jumlah cluster atau pengelompokan sebanyak 3 kelompok. Pengelompokan 1 antara titik pusat *centroid* dan anggota *cluster* memiliki jarak yang terdekat, sehingga karakteristik penjualan asesoris dan *sparepart* komputer pada toko laksamana komputer adalah produk *vga splitter 4 port* dan *cd blank/pack* merk *tdk* untuk pengelompokan 2 adalah produk *handset gaming*, *Converter hdmi to vga*, *vga splitter 2 port*, tinta canon 1 liter colour dan kabel usb 1 pack, untuk pengelompokan 3 adalah produk *mouse wereless*, *cd blank/pack* merk *gt pro*, kabel *vga 15 m*, kabel *hdmi 10m* dan kabel *usb 1 pack*. Maka dapat ditarik kesimpulan bahwa jika konsumen melakukan pembelian asesoris dan *sparepart* ditoko laksamana komputer dumai maka pihak toko akan memberikan posisi produk berdasarkan banyaknya minat konsumen terhadap produk tersebut.

**4. SIMPULAN**

Dalam penelitian ini dapat dianalisa dan disimpulkan beberapa hal sebagai berikut:

1. Dengan analisis berdasarkan produk yang sering laris terjual data yang diperoleh toko Laksamana Komputer Dumai dapat menghasilkan sebuah informasi mengenai data prediksi pengelompokan minat konsumen terhadap *sparepart* dan asesoriskomputer.
2. Pengelompokan data ke dalam cluster secara iterative dari penjualan *sparepart* dan asesoris komputer di Toko Laksamana Komputer Dumai dari hasil analisa data mining adalah pada cluster 2 dan cluster 3 dan produk yang sering diminati konsumen akan diposisikan paling depan sertadi kelompokkan pada produk yang sering diminati.
3. Dengan menggunakan *Algoritma K-Means Clustering* maka dapat menentukan arketipe pembelian asesoris dan *sparepart* ditoko Laksamana Komputer Dumai.

**REFERENSI**

- Aras, Z., & Sarjono. (2016). Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means (Studi Kasus : Kantor kecamatan Bahar Utara). *Jurnal Manajemen Sistem Informasi* , ISSN : 2540-8011, Hal:159-170.
- Dahlan, A., Utami, E., & Luthfi, E. T. (2013). Perancangan Data Warehouse Perpustakaan Perguruan Tinggi XYZ Menggunakan Metode Snowflake Schema. *Teknologi Informasi* , ISSN:1907-2430 Hal:1-15.
- Fahmi, M. (2014). Klasterisasi Judul Buku Dengan Menggunakan Metode K-Means. *SNATI* , ISSN: 1907-5022 Hal:1-4.
- Jumaroh, S., Widodo, N., & Hirawan, W. (2015). Analisa Klaster K-Means Dari Data Luas Grup Sunspot Dan Data Grup Sunspot Klasifikasi MC.Intosh Yang Membangkitkan Flare Soft X-Ray Dan H. *Cauci: Jurnal Matematika dan Aplikasi* , ISSN:2086-0382 Hal: 1-9.
- Junanto, A. (2013). Algoritma Naive Bayes Untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Jurnal Teknologi Informadi DINAMK* , ISSN : 0854 : 9524 Hal: 9-16.
- M.Agustin, F. E., Fitria, A., & S, A. H. (2015). Implementasi Algoritma K-means Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus : SMP NEGERI 101 JAKARTA). *Jurnal Teknik Informatika* , ISSN : 2549-0923 Hal : 73-78.
- Madhulatha, T. S. (2012). An Overview On Clustering Methods. *Jurnal Of Engineering* , ISSN:2250-3021 Hal:1-7.
- Mulyati, S. (2015). Penerapan Data Mining Dengan Metode Clustering Untuk Pengelompokan Data Pengiriman Burung. *Seminar Ilmiah Nasional Teknologi Komputer* , ISSN:2460-4690 Hal:30-35.
- Nasari, F., & Darma, S. (2015). Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama). *Seminar Nasiona Teknologi Informasi Dan Multimedia* , ISSN; 2302-3805 Hal:1-8.
- Nasari, F., & Darma, S. (2015). Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru. *Seminar Nasional Teknologi Informasi* , ISSN : 2302-3805 Hal: 73-78.
- Ndaumanu, R. I., Kusriani, & Arief, M. R. (2014, September). Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa Dengan Metode K-Nearest Neighbor. *Jurnal Teknik Informatika Dan Sistem Informasi* , ISSN : 2407-4322 Hal: 2-5.
- Nurjoko, & Darmawan, A. (2015, Mei). Penerapan Data Mining Menggunakan Association Rules Untuk Mendukung Strategi Pemasaran Calon Mahasiswa Baru. *Tim Darmajaya* , ISSN : 2442-556 hal:17-19.
- Ong, J. O. (2013). Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University. *Jurnal Ilmiah Teknik Industri* , ISSN:1412-6869 Hal:10-19.
- Ramadhani, A., Farmadi, A., & Budiman, I. (1 Juni 2014). Clustering Data Cuaca Untuk Pengenalan Pola Prioditas Iklim Wilayah Pelaihari Dengan Metode Fuzzy C-Means. *Teknik Dan Industri* , ISSN : 2087-6920 Hal: 57-64.
- Sari, C. R. (2016). Teknik Data Mining Menggunakan Classification Dalam Sistem Penunjang Keputusan Peminatan SMA Negeri 1 Polewali. *IJNS* , ISSN : 1979-9330 Hal:48-54.
- Sunjana. (2010). Aplikasi Mining Data Mahasiswa Dengan Metode Klasifikasi Decision Tree. *Seminar Nasional Aplikasi Teknologi Informasi* , ISSN : 1907-5022, Hal:24-29.