



Artikel

## IMPLEMENTASI METODE CLUSTERING K-MEANS UNTUK REKOMENDASI PENGADAAN STOK LAMPU DI PT GLOBAL LIGHTING INDONESIA

Prayogo Putra Tjaya<sup>1</sup>, Rino<sup>2</sup>, Hartana Wijaya<sup>3</sup>

<sup>1, 2, 3</sup> Universitas Buddhi Dharma, Teknik Informatika, Banten, Indonesia

### SUBMISSION TRACK

Received: August 25, 2021

Final Revision: September 25, 2021

Available Online: September, 2021

### KATA KUNCI

Stock Lampu, Clustering, K-Mean, DataMining

### KORESPONDENSI

Phone: 08179091499

E-mail: prayogo.p.t@gmail.com

### ABSTRAK

PT Global Lighting Indonesia salah satu perusahaan lampu yang sedang berkembang di Indonesia. Dalam suatu usaha perdagangan, jumlah permintaan dari konsumen yang tidak menentu mengakibatkan stok yang disiapkan menjadi tidak stabil. Pengadaan stok lampu di PT Global Lighting Indonesia hanya berdasarkan perkiraan. Hal ini disebabkan karena pemilik perusahaan tidak ingin terjadinya kekurangan stok lampu pada saat permintaan konsumen dalam jumlah besar, maka perusahaan mengambil langkah yaitu melakukan pemesanan barang lebih besar dari sebelumnya. Berdasarkan permasalahan diatas maka di butuhkan sistem dan metode yang mampu mengelompokan produk berdasarkan merk atau model tertentu sehingga memudahkan saat pembuatan laporan dan pengadaan produk, sehingga bisa lebih mengefisiensi waktu, tenaga serta biaya. Juga untuk menghindari resiko pengambilan stok lampu yang kurang laku terambil kembali, sehingga berpengaruh terhadap tinggi rendahnya angka penjualan di perusahaan dan terjadinya penumpukan stok lampu yang mengakibatkan barang menjadi rusak dan tidak dapat terjual lagi. Pada penelitian ini, peneliti akan membuat sebuah aplikasi menggunakan metode *Clustering K-means* untuk mengolah data transaksi yang masuk sehingga akan memudahkan pemilik dan karyawan PT Global Lighting Indonesia dalam manajemen stok lampu.

### PENGANTAR

PT Global Lighting Indonesia salah satu perusahaan lampu yang sedang berkembang di Indonesia. Dalam suatu usaha perdagangan, jumlah permintaan dari konsumen yang tidak menentu mengakibatkan stok yang disiapkan menjadi tidak stabil. Pengadaan stok lampu di PT Global Lighting Indonesia hanya berdasarkan perkiraan. Hal ini disebabkan karena pemilik perusahaan tidak ingin

terjadinya kekurangan stok lampu pada saat permintaan konsumen dalam jumlah besar, maka perusahaan mengambil langkah yaitu melakukan pemesanan barang lebih besar dari sebelumnya. Hal ini dapat mengatasi kekurangan persediaan stok lampu dan perusahaan tidak perlu melakukan pemesanan berulang-ulang kedistributor sehingga dapat berpengaruh terhadap menurunnya penjualan di perusahaan, selain itu manajemen stok yang tidak akurat juga

mengakibatkan sering terjadi kekurangan produk yang dapat merugikan perusahaan dan juga mengecewakan konsumen. Metode *K-Means* merupakan metode *clustering* yang paling sederhana dan umum. *K-Means* merupakan salah satu algoritma *clustering* dengan metode partisi (*partitioning method*) yang berbasis titik pusat (*centroid*) selain algoritma *k-Medoids* yang berbasis obyek. Algoritma ini pertama kali diusulkan oleh MacQueen (1967) dan dikembangkan oleh Hartigan dan Wong tahun 1975 dengan tujuan untuk dapat membagi  $M$  data point dalam  $N$  dimensi kedalam sejumlah  $k$  *cluster* dimana proses *clustering* dilakukan dengan meminimalkan jarak *sum squares* antara data dengan masing masing pusat (*centroid-based*). Pada penelitian ini, peneliti akan membuat sebuah aplikasi menggunakan metode *Clustering K-means* untuk mengolah data transaksi yang masuk sehingga akan memudahkan pemilik dan karyawan PT Global Lighting Indonesia dalam manajemen stok lampu. Penelitian ini diharapkan dapat menghasilkan suatu model program yang dapat *mengcluster* atau mengelompokkan produk yang memiliki jumlah transaksi penjualan banyak karena paling diminati, jumlah transaksi penjualan sedang karena produk diminati, dan jumlah transaksi penjualan dikit karena kurang diminati dari beragam produk yang banyak. Berdasarkan uraian latar belakang tersebut, maka penelitian ini di beri judul **“Implementasi Metode *Clustering K-means* Untuk Rekomendasi Pengadaan Stock Lampu Pada PT Global Lighting Indonesia”**.

## I. METODE

### 1.1. Data Mining

Kata *Mining* merupakan kiasan dari bahasa inggris, *mine*. Jika *mine* berarti menambang sumber daya yang tersembunyi di dalam tanah, maka *Data Mining* merupakan penggalian makna yang tersembunyi dari kumpulan data yang sangat besar. Karena itu *Data Mining* sebenarnya memiliki akar yang

panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan basis Data[1].

*Data Mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Data Mining* adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar[2]. Berdasarkan definisi diatas, data mining dapat disimpulkan sebagai serangkaian proses yang berkaitan dengan pencarian pola pada suatu data yang berfungsi untuk menggali nilai tambah dari data dengan menggunakan teknologi pengenalan pola, statistik, dan matematika

### 1.2. Clustering

*Clustering* merupakan proses pengelompokan sekumpulan objek ke dalam beberapa kelas. Tujuannya adalah mengelompokkan objek data menjadi *cluster* terpisah sehingga tiap objek pada *cluster* yang sama memiliki kemiripan karakteristik, sedangkan objek yang memiliki karakteristik berbeda akan ditempatkan pada *cluster* yang lain. Objek-objek dalam suatu *cluster* yang sama harus semirip mungkin satu sama lain dan objek-objek dalam suatu *cluster* harus berbeda dengan objek-objek dalam *cluster* lain. Cara kerja data *clustering* adalah menempatkan data-data yang ada ke dalam beberapa kelas yang mempunyai kesamaan (kemiripan). *Cluster* merupakan pola yang terbentuk dalam suatu proses pembagian sekelompok data ke dalam sejumlah *sub-kelas*. *Clustering* dapat digunakan untuk memberikan label suatu data yang belum diketahui kelasnya. Pada dasarnya terdapat dua metode dalam *clustering* yaitu *Hierarchical Clustering Method* dan *Non Hierarchical Clustering Method*. *Hierarchical Cluster Method* digunakan apabila belum ada informasi jumlah *cluster*. Prinsip utamanya adalah mengatur seluruh

objek ke dalam sebuah pohon berdasarkan suatu kriteria tertentu. Pohon itu disebut dengan *dendogram*. Semakin tinggi level dari simpul pohon, maka semakin rendah similaritas antar objeknya. Objek yang memiliki kesamaan berada pada kelompok awal atau simpul pohon paling bawah. Sedangkan *non hierarchical clustering method* pertemuan mengelompokkan objek ke dalam keadaan yang dapat dikatakan jumlah kelompok pada metode diturunkan oleh pengguna[4]. Terdapat dua algoritma dalam *non hierarchical clustering method* :

#### 1. Single Pass

Algoritma ini melewati data dari suatu kumpulan data sekali saja untuk menciptakan *cluster*.

#### 2. K-means

Melewati data berulang kali untuk mengadakan pengujian terhadap pusat *cluster* dan menentukan pengelompokan terbaik dengan mendefinisikan jumlah *cluster* terlebih dahulu.

Pengelompokan data menjadi sejumlah kategori juga dapat dilakukan menggunakan metode *clustering*. Berbeda dengan metode klasifikasi, *clustering* mengelompokkan data hanya berdasarkan fitur-fitur yang ada pada data tersebut. Berdasarkan sifat tersebut, *clustering* tidak memerlukan data latih yang telah diketahui kelasnya. Oleh karena itu, proses pembelajaran pada *clustering* bersifat mandiri, yang sering disebut dengan istilah *unsupervised learning*. Tujuan utama dari proses *clustering* adalah membagi sekumpulan data menjadi sekumpulan grup (*cluster*) sehingga data-data dalam satu *cluster* memiliki banyak kemiripan namun berbeda dengan data-data yang berada pada *cluster* lainnya. Kemiripan antar data dapat dihitung menggunakan berbagai metode pengukuran jarak yang telah dijelaskan pada bab sebelumnya. Teknik *clustering* memiliki banyak kegunaan dalam kehidupan sehari-hari. Bidang biologi memanfaatkan *clustering* untuk membentuk taksonomi dari

berbagai tumbuhan. Taksonomi merupakan pengelompokan makhluk hidup menjadi tingkatan-tingkatan kelompok tertentu. *Clustering* juga dapat digunakan untuk mengelompokkan daerah-daerah dengan kondisi cuaca yang mirip. Dengan mengetahui pola cuaca di suatu daerah, pemilihan tanaman yang cocok dibudidayakan di daerah tersebut dapat dilakukan dengan lebih akurat. Mesin pencarian informasi, seperti *google* dan *yahoo*, dapat mengelompokkan hasil pencarian sehingga pengguna dapat menelusuri hasil pencarian dari kelompok yang relevan saja. Hal ini dapat mengurangi waktu pencarian dan memperbesar peluang mendapatkan hasil yang sesuai.

### 1.3. Algoritma K-Means

[7] Metode *K-Means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien. *K-Means* merupakan salah satu algoritma *clustering* dengan metode partisi (*partitioning method*) yang berbasis titik pusat (*centroid*) selain algoritma *k-Medoids* yang berbasis obyek. Algoritma ini pertama kali diusulkan oleh MacQueen (1967) dan dikembangkan oleh Hartigan dan Wong tahun 1975 dengan tujuan untuk dapat membagi  $M$  data point dalam  $N$  dimensi kedalam sejumlah  $k$  *cluster* dimana proses *clustering* dilakukan dengan meminimalkan jarak *sum squares* antara data dengan masing masing pusat (*centroid-based*). Algoritma dalam penerapannya memerlukan tiga parameter yang seluruhnya ditentukan pengguna yaitu jumlah *cluster*  $k$ , inisialisasi klaster, dan jarak *system*. Biasanya, *K-Means* dijalankan secara independen dengan inisialisasi yang berbeda menghasilkan *cluster* akhir yang berbeda karena algoritma ini secara prinsip hanya mengelompokkan data menuju *local minimal*. Salah satu cara untuk mengatasi *local minimal* adalah dengan mengimplementasikan algoritma *K-Means*, untuk  $K$  yang

diberikan, dengan beberapa nilai initial partisi yang berbeda dan selanjutnya dipilih partisi dengan kesalahan kuadrat terkecil[5]. Berikut adalah langkah-langkah algoritma *K- Means* :

1. Tentukan berapa banyak *cluster* k dari dataset yang akan dibagi.
2. Tetapkan secara acak data k menjadi pusat awal lokasi klaster.
3. Untuk masing-masing data, temukan pusat *cluster* terdekat. Dengan demikian berarti masing-masing pusat *cluster* memiliki sebuah *subset* dari *dataset*, sehingga mewakili bagian dari *dataset*. Oleh karena itu, telah terbentuk *cluster* k: C1, C2, C3, ..., Ck.
4. Untuk masing-masing *cluster* k, temukan pusat luasan klaster, dan perbarui lokasi dari masing masing pusat *cluster* ke nilai baru dari pusat luasan.
5. Ulangi langkah ke-3 dan ke-5 hingga data-data pada tiap *cluster* menjadi terpusat atau selesai.

### 1.5. *Cross industry standard process for data mining (CRISP-DM)*

CRISP-DM (*cross-industry standard process for data mining*) merupakan suatu konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam *data mining* yang dapat diaplikasikan di berbagai sektor industri. Berikut ini adalah enam tahap siklus hidup pengembangan *data mining* :

#### 1. *Business Understanding*

Tahap pertama adalah memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemahkan pengetahuan ini ke dalam pendefinisian masalah pada data mining. Selanjutnya akan ditentukan rencana dan strategi untuk mencapai tujuan tersebut.

#### 2. *Data Understanding*

Tahap ini dimulai dengan pengumpulan data yang kemudian akan dilanjutkan

dengan proses untuk mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa untuk informasi yang tersembunyi.

#### 3. *Data Preparation*

Tahap ini meliputi semua kegiatan untuk membangun *dataset* akhir (data yang akan diproses pada tahap pemodelan) dari data mentah. Tahap ini dapat diulang beberapa kali. Pada tahap ini juga mencakup pemilihan tabel, *record*, dan atribut- atribut data, termasuk proses pembersihan dan transformasi data untuk kemudian dijadikan masukan dalam tahap pemodelan.

#### 4. *Modelling*

Dalam tahap ini akan dilakukan pemilihan dan penerapan berbagai teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal. Secara khusus, ada beberapa teknik berbeda yang dapat diterapkan untuk masalah *data mining* yang sama. Di pihak lain ada teknik pemodelan yang membutuhkan format data khusus. Sehingga pada tahap ini masih memungkinkan kembali ke tahap sebelumnya.

#### 5. *Evaluation*

Pada tahap ini, model sudah terbentuk dan diharapkan memiliki kualitas baik jika dilihat dari sudut pandang analisa data. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas model sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang ditetapkan pada fase awal (*Business Understanding*). Kunci dari tahap ini adalah menentukan apakah ada masalah bisnis yang belum dipertimbangkan. Di akhir dari tahap ini harus ditentukan penggunaan hasil proses *data mining*.

6. *Deployment*

Pada tahap ini, pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna.

Tahap *deployment* dapat berupa pembuatan laporan sederhana atau mengimplementasikan proses *data mining* yang berulang dalam perusahaan. Pada banyak kasus, tahap *deployment* melibatkan konsumen, di samping analisis data, karena sangat penting bagi konsumen untuk memahami tindakan apa yang harus dilakukan untuk menggunakan model yang telah dibuat.

1.6. *Normalization*

Normalization adalah proses pengelompokan atribut ke dalam hubungan yang terstruktur dengan baik dan bebas dari anomali. Normalization digunakan untuk mentransformasi sebuah atribut numerik diskalakan dalam range yang lebih kecil seperti -1.0 sampai 1.0. Teknik yang dapat digunakan untuk operasi ini adalah min-max normalization, Z-Score Normalization dan normalization by decimal scaling[9].

1. Min-max Normalization

Min-max normalization memetakan sebuah value  $v$  dari atribut  $A$  menjadi  $v'$  ke dalam range  $[new\_min_A, new\_max_A]$  berdasarkan rumus:

$$v^1 = \frac{v - min_A}{max_A - min_A} (new\_max_A - new\_min_A) + new\_min_A$$

2. Z-Score Normalization

Disebut juga zero-mean normalization, dimana value dari sebuah atribut  $A$  dinormalisasi berdasarkan nilai rata-rata dan standar deviasi dari atribut  $A$ . Sebuah value  $v$  dari atribut  $A$  dinormalisasi menjadi  $v'$  dengan rumus: ( $\mu$ : mean,  $\sigma$ : standard deviation).

$$v^1 = \frac{v - \mu_A}{\sigma_A}$$

3. Normalization by Decimal Scaling

Normalisasi yang diperoleh dengan melakukan penggeseran titik desimal dari value sebuah atribut  $A$ . Jumlah titik desimal yang digeser tergantung dari nilai absolut maksimum dari atribut  $A$ .

$$v^1 = \frac{v}{10^j}$$

Dimana  $j$  adalah integer terkecil sehingga  $Max(|v^1|) < 1$ .

II. PERANCANGAN

2.1. *Fase Evaluation*

Pada Tahap ini *dataset* yang sudah disiapkan kemudian akan diolah menggunakan aplikasi *Rapidminer* 9.8. Tahap ini dimulai menyiapkan data sampel yang sudah ditentukan untuk diolah dan dianalisis. Evaluasi disini akan dilakukan secara mendalam dengan tujuan untuk menyesuaikan hasil pada tahap *modeling* agar sesuai dengan sasaran yang ingin dicapai dalam tahap *Business Understanding*. Hasil evaluasi adalah tahap dimana menemukan *centroid* yang dapat digunakan sebagai sumber informasi baru, menemukan *centroid* dari data transaksi yang kuat yang memenuhi persyaratan yang ada. Pada data *itemset* yang sudah disediakan akan diunggah kedalam *Rapidminer* sebagai langkah awal, selanjutnya memilih algoritma yang akan digunakan yaitu *K-means*.

2.2. *Fase Penyebaran (Fase Deployment)*

Fase *Deployment* merupakan tahapan membuat laporan hasil kegiatan *data mining*. Laporan akhir mengenai pengetahuan yang didapat atau pengenalan pola pada data dalam proses *data mining* dan dipresentasikan dalam bentuk grafik atau deskripsi yang mudah dipahami. Dan berikut pengetahuan atau hasil dari proses *data mining* menggunakan *rapidminer* 9.8 :

**Cluster Model**

Cluster 0: 168 items  
 Cluster 1: 5 items  
 Cluster 2: 19 items  
 Total number of items: 192

**III. PEMBAHASAN**

**3.1. Perhitungan Manual**

Data testing yang digunakan untuk proses

perhitungan manual menggunakan data sampel sebanyak 192 data sampel yang di ambil dari data transaksi yang didapat dari penjualan PT. Global Lighting Indonesia.

**Tabel1. Data Transaksi**

| Nama Barang                                    | Jumlah Unit Terjual | Total Banyak Transaksi |
|--|---------------------|------------------------|
| LAMP HOUSING CVL 001 MT.CHROME 3000K CIVILITE  | 40                  | 2                      |
| LAMP HOUSING CVL 001 FWH WHITE 3000K CIVILITE  | 100                 | 1                      |
| LED SPOT LIGHT CVL-005 5W 3000K BLACK CIVILITE | 420                 | 1                      |
| LED SPOT LIGHT CVL-005 5W 6000K BLACK CIVILITE | 385                 | 6                      |
| LED SPOT LIGHT CVL-005 5W 3000K WHITE CIVILITE | 660                 | 8                      |

Dari Tabel Data Transaksi di transformasikan menggunakan rumus *min- max normalization*. Rumus *min-max normalization* :

$$v^1 = \frac{v - \min_A}{\max_A - \min_A} (\text{new\_max} - \text{new\_min}) + \text{new\_min}$$

1. Menghitung nilai pada kolom jumlah unit terjual, yaitu :  
 Diketahui :  
 Nilai Min = 1 dan Max = 8000  
 New Min = 1 dan New Max = 100  
 Perhitungan menggunakan rumus *min-max* :

$$v^1 = \frac{40 - 1}{800 - 1} (100 - 1) + 1 = 1,48$$

2. Menghitung nilai pada kolom jumlah unit terjual, yaitu :

Diketahui :  
 Nilai Min = 1 dan Max = 42  
 New Min = 1 dan New Max = 100  
 Perhitungan menggunakan rumus *min-max* :

$$v^1 = \frac{2 - 1}{42 - 1} (100 - 1) + 1 = 3,41$$

**Tabel2. Hasil Perhitungan Min-Max Normalization**

| Nama Barang                                    | Jumlah Unit Terjual | Total Banyak Transaksi |
|--|---------------------|------------------------|
| LAMP HOUSING CVL 001 MT.CHROME 3000K CIVILITE  | 1,48                | 3,41                   |
| LAMP HOUSING CVL 001 FWH WHITE 3000K CIVILITE  | 2,23                | 1,00                   |
| LED SPOT LIGHT CVL-005 5W 3000K BLACK CIVILITE | 6,19                | 8,24                   |
| LED SPOT LIGHT CVL-005 5W 6000K BLACK CIVILITE | 5,75                | 13,07                  |
| LED SPOT LIGHT CVL-005 5W 3000K WHITE CIVILITE | 9,16                | 17,90                  |

Tahap Selanjutnya adalah menentukan centroid dengan memilih secara acak 3 titik *centroid* yang ada pada data transaksi.

**Tabel3. Titik Pusat Awal**

| Centroid | Jumlah Unit Terjual (X) | Total Banyak Transaksi (Y) |
|----------|-------------------------|----------------------------|
| C1       | 2,10                    | 5,83                       |
| C2       | 1,07                    | 3,41                       |
| C3       | 14,29                   | 15,49                      |

Setelah mengetahui nilai K dan titik pusat *cluster* awal, kemudian mengukur jarak antara pusat cluster menggunakan rumus *euclidian distance*. Dan akan di dapatkan matriks jarak C1,C2, dan C3 sebagai berikut :

Rumus *euclidian distance* :

$$[(x, y), (a, b)] \sqrt{(x - a)^2 + (y - b)^2}$$

1. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-1, yaitu :

$$d(1,1) = \sqrt{(1,48 - 2,10)^2 + (3,41 - 5,83)^2} = 2,49$$

2. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-2, yaitu :

$$d(1,2) = \sqrt{(1,48 - 1,07)^2 + (3,41 - 3,41)^2} = 0,41$$

3. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-3, yaitu :

$$d(1,3) = \sqrt{(1,48 - 14,29)^2 + (3,41 - 15,48)^2} = 17,6$$

**Tabel4. Hasil Perhitungan Jarak Pusat Cluster Iterasi Ke 1**

| Data Ke | C1    | C2    | C3    | Cluster |
|---------|-------|-------|-------|---------|
| 1       | 2,49  | 0,41  | 17,60 | C2      |
| 2       | 4,83  | 2,67  | 18,86 | C2      |
| 3       | 4,47  | 7,03  | 10,87 | C1      |
| 4       | 8,11  | 10,73 | 8,87  | C1      |
| 5       | 13,98 | 16,59 | 5,68  | C3      |

Dari hasil perhitungan jarak akan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam kelompok yang sudah ditentukan. Dan berikut ini merupakan tampilan tabel pengelompokan data, nilai 0 (kosong) merupakan data tersebut tidak berada dalam kelompok data dan sedangkan yang berisi nilai berarti data tersebut berada dalam kelompok data.

**Tabel5. Hasil Pengelompokan Data**

| Rerata | Cluster 1 |       | Cluster 2 |      | Cluster 3 |      |
|--------|-----------|-------|-----------|------|-----------|------|
|        | X         | Y     | X         | Y    | X         | Y    |
| 1      | 0,00      | 0,00  | 1,48      | 3,41 | 0,00      | 0,00 |
| 2      | 0,00      | 0,00  | 2,23      | 1,00 | 0,00      | 0,00 |
| 3      | 6,19      | 8,24  | 0,00      | 0,00 | 0,00      | 0,00 |
| 4      | 5,75      | 13,07 | 0,00      | 0,00 | 0,00      | 0,00 |
| 5      | 0,00      | 0,00  | 0,00      | 0,00 | 9,16      | 17,9 |

Kemudian setelah mengetahui anggota setiap *cluster* adalah mencari centroid baru berdasarkan rata-rata nilai tiap *cluster*. *Cluster* baru tersebut didapat dari rumus = nilai hasil/banyak hasil. Berikut perhitungannya :

$$\text{Cluster 1}(x) = 183,05 / 65 = 2,28$$

$$\text{Cluster 1}(y) = 610,71 / 65 = 9,40$$

$$\text{Cluster 2}(x) = 135,64 / 87 = 1,56$$

$$\text{Cluster 2}(y) = 154,61 / 87 = 1,78$$

$$\text{Cluster 3}(x) = 667,53 / 40 = 16,69$$

$$\text{Cluster 3}(y) = 1225,59 / 40 = 30,64$$

**Tabel6. Titik Pusat Cluster Baru**

| Centroid | Jumlah Unit Terjual (X) | Total Banyak Transaksi (Y) |
|----------|-------------------------|----------------------------|
| C1       | 2,82                    | 9,40                       |
| C2       | 1,56                    | 1,78                       |
| C3       | 16,69                   | 30,64                      |

Ulangi perhitungan jarak data pertama dengan pusat *cluster* baru.

1. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-1, yaitu :

$$d(1,1) = \sqrt{(1,48 - 2,82)^2 + (3,41 - 9,40)^2} = 6,13$$

2. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-2, yaitu :

$$d(1,2) = \sqrt{(1,48 - 1,56)^2 + (3,41 - 5,78)^2} = 1,64$$

3. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-3, yaitu :

$$d(1,3) = \sqrt{(1,48 - 16,69)^2 + (3,41 - 30,64)^2} = 31,18$$

**Tabel7. Hasil Perhitungan Jarak Pusat Cluster Iterasi Ke 2**

| Data Ke | C1    | C2    | C3    | Cluster |
|---------|-------|-------|-------|---------|
| 1       | 6,13  | 1,64  | 31,18 | C2      |
| 2       | 8,42  | 1,02  | 32,98 | C2      |
| 3       | 3,56  | 7,95  | 24,74 | C1      |
| 4       | 4,71  | 12,05 | 20,69 | C1      |
| 5       | 10,61 | 17,83 | 28,78 | C1      |

Kemudian selanjutnya jarak hasil perhitungan akan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa pada data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.

**Tabel8. Hasil Pengelompokan Data**

| Rerata  | Cluster 1 |       | Cluster 2 |      | Cluster 3 |      |
|---------|-----------|-------|-----------|------|-----------|------|
| Data Ke | X         | Y     | X         | Y    | X         | Y    |
| 1       | 0,00      | 0,00  | 1,48      | 3,41 | 0,00      | 0,00 |
| 2       | 0,00      | 0,00  | 2,23      | 1,00 | 0,00      | 0,00 |
| 3       | 6,19      | 8,24  | 0,00      | 0,00 | 0,00      | 0,00 |
| 4       | 5,75      | 13,07 | 0,00      | 0,00 | 0,00      | 0,00 |
| 5       | 9,16      | 17,90 | 0,00      | 0,00 | 0,00      | 0,00 |

Kemudian setelah mengetahui anggota setiap *cluster* adalah mencari centroid baru berdasarkan rata-rata nilai tiap *cluster*. *Cluster* baru tersebut didapat dari rumus = nilai hasil/banyak hasil. Berikut perhitungannya :

$$\text{Cluster 1}(x) = 289,92 / 79 = 2,28$$

$$\text{Cluster 1}(y) = 895,15 / 79 = 6,67$$

$$\text{Cluster 2}(x) = 174,24 / 91 = 1,91$$

$$\text{Cluster 2}(y) = 165,85 / 91 = 1,82$$

$$\text{Cluster 3}(x) = 522,05 / 22 = 23,73$$

$$\text{Cluster 3}(y) = 929,90 / 22 = 42,27$$

**Tabel9. Titik Pusat Cluster Baru**

| Centroid | Jumlah Unit Terjual (X) | Total Banyak Transaksi (Y) |
|----------|-------------------------|----------------------------|
| C1       | 3,67                    | 11,33                      |
| C2       | 1,91                    | 1,82                       |
| C3       | 23,73                   | 42,27                      |

Ulangi perhitungan jarak data pertama dengan pusat *cluster* baru.

1. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-1, yaitu :

$$d(1,1) = \sqrt{(1,48 - 3,67)^2 + (3,41 - 11,33)^2} = 8,21$$

2. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-2, yaitu :

$$d(1,2) = \sqrt{(1,48 - 1,91)^2 + (3,41 - 1,82)^2} = 1,65$$

3. Menghitung jarak data Ke-1 dengan pusat *cluster* Ke-3, yaitu :

$$d(1,3) = \sqrt{(1,48 - 23,73)^2 + (3,41 - 42,27)^2} = 44,77$$

**Tabel10. Hasil Perhitungan Jarak Pusat Cluster**

| Iterasi Ke 3 |       |       |       |         |
|--------------|-------|-------|-------|---------|
| Data Ke      | C1    | C2    | C3    | Cluster |
| 1            | 8,21  | 1,65  | 44,77 | C2      |
| 2            | 10,43 | 0,88  | 46,53 | C2      |
| 3            | 3,98  | 7,71  | 38,28 | C1      |
| 4            | 2,72  | 11,89 | 34,29 | C1      |
| 5            | 8,56  | 17,64 | 28,39 | C1      |

Kemudian selanjutnya jarak hasil perhitungan akan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa pada data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.

**Tabel11. Hasil Pengelompokan Data**

| Rerata  | Cluster 1 |       | Cluster 2 |      | Cluster 3 |      |
|---------|-----------|-------|-----------|------|-----------|------|
| Data Ke | X         | Y     | X         | Y    | X         | Y    |
| 1       | 0,00      | 0,00  | 1,48      | 3,41 | 0,00      | 0,00 |
| 2       | 0,00      | 0,00  | 2,23      | 1,00 | 0,00      | 0,00 |
| 3       | 6,19      | 8,24  | 0,00      | 0,00 | 0,00      | 0,00 |
| 4       | 5,75      | 13,07 | 0,00      | 0,00 | 0,00      | 0,00 |
| 5       | 9,16      | 17,90 | 0,00      | 0,00 | 0,00      | 0,00 |

Kemudian setelah mengetahui anggota setiap *cluster* adalah mencari centroid baru berdasarkan rata-rata nilai tiap *cluster*. *Cluster* baru tersebut didapat dari rumus = nilai hasil/banyak hasil. Berikut perhitungannya :

$$\text{Cluster 1}(x) = 320,69 / 69 = 4,65$$

$$\text{Cluster 1}(y) = 955,17 / 69 = 13,84$$

$$\text{Cluster 2}(x) = 208,06 / 108 = 1,93$$

$$\text{Cluster 2}(y) = 264,95 / 108 = 2,45$$

$$\text{Cluster 3}(x) = 457,46 / 15 = 30,50$$

$$\text{Cluster 3}(y) = 770,78 / 15 = 51,39$$

Pada perhitungan ini berhenti pada iterasi ke-11 dimana kelompok data ke-11 sama dengan kelompok data ke-10 tidak berubah, maka hasil sudah sesuai dengan pengelompokan *cluster*. Dan user dapat menentukan hasil pengelompokan data masuk kedalam kategori Paling Laku, Laku,

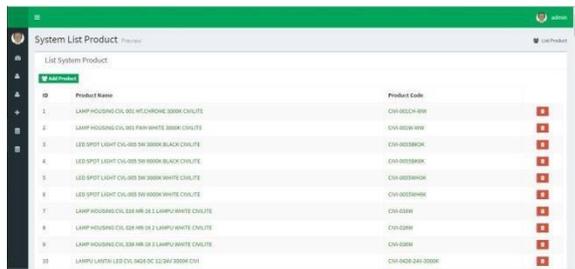
dan Kurang Laku, Berikut ini adalah hasil dari perhitungan manual tersebut :

**Tabel11. Hasil Perhitungan Manual**

| Data Ke | Jumlah Unit Terjual | Total Banyak Transaksi | Cluster | Hasil (User) |
|---------|---------------------|------------------------|---------|--------------|
| 1       | 1,48                | 3,41                   | C2      | KL           |
| 2       | 2,23                | 1,00                   | C2      | KL           |
| 3       | 6,19                | 8,24                   | C2      | KL           |
| 4       | 5,75                | 13,07                  | C2      | KL           |
| 5       | 9,16                | 17,90                  | C1      | L            |

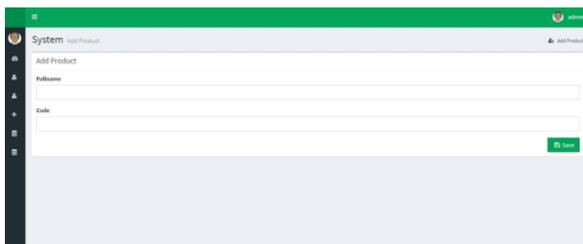
**3.2. Tampilan Program**

Berikut ini adalah tampilan dari aplikasi berbasis web yang telah dihasilkan :



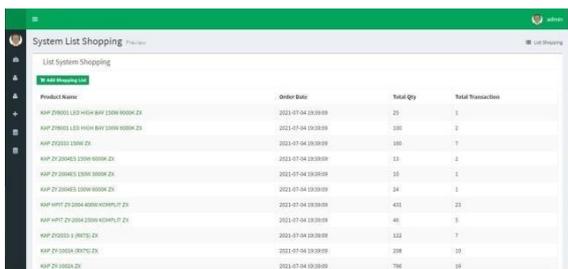
**Gambar 4.9 Halaman List Product**

Halaman list product adalah halaman untuk melihat dan mengubah daftar barang yang tersedia.



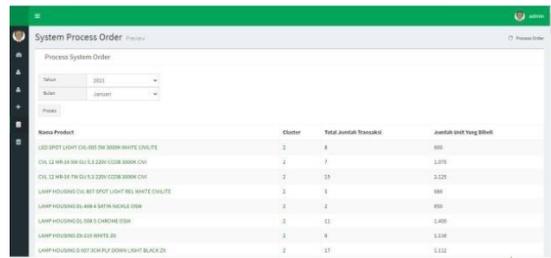
**Gambar 4.10 Halaman Add Product**

Halaman add product adalah halaman untuk menambahkan barang baru.



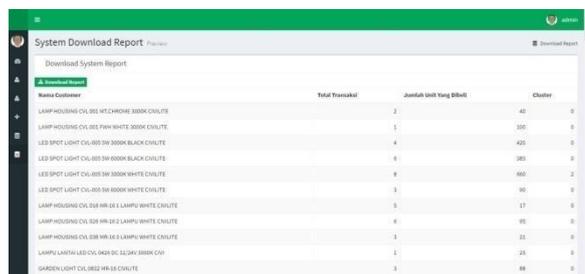
**Gambar 4.11 Halaman List Shopping**

Halaman list shopping adalah halaman untuk melihat dan menambah transaksi.



**Gambar 4.12 Halaman Process Order**

Halaman process order adalah halaman untuk membuat rekomendasi pengadaan barang menggunakan algoritma K-Means.



**Gambar 4.13 Halaman Download Report**

Halaman download report adalah halaman untuk mengunduh hasil rekomendasi pengadaan barang menggunakan algoritma K-Means.

**IV. KESIMPULAN**

Berdasarkan hasil yang didapat dari penulisan skripsi ini, maka diperoleh kesimpulan sebagai berikut :

1. Dengan adanya pengelompokan data ini, pihak perusahaan dapat mengetahui barang Paling Laku, Laku, dan Kurang Laku. Sehingga dapat memberikan solusi kepada pihak perusahaan dalam menentukan stok barang.
2. Dalam penelitian ini *output* yang dihasilkan dari 3 *Cluster* yaitu, *Cluster* 0 dengan 39 data, *Cluster* 1 dengan 146 data, dan *Cluster* 2 dengan 7 data dengan menggunakan 192 sampel *dataset*. Dari simpulan yang dihasilkan peneliti mengkategorikan *Cluster* 0 adalah Laku, *Cluster* 1 adalah Kurang Laku, dan *Cluster* 2 adalah Paling Laku.

## DAFTAR PUSTAKA

- [1]. C, Aprilla Dennis., dkk. 2013. Belajar Data Mining Dengan Rapid Miner. Jakarta.
- [2]. Nofriansyah, D. 2015. Konsep Data Mining Vs Sistem Pendukung Keputusan. Yogyakarta : Deepublish.
- [3]. Buulolo, Efori. 2020. Data Mining Untuk Perguruan Tinggi. Yogyakarta : Deepublish.
- [4]. Purwanto, Ari., dan Samingun Handoyo. 2017. Sistem Fuzzy Terapan Dengan Software R. Malang : UB Press.
- [5]. Adinugroho, Sigit., dan Yuita Arum Sari. 2018. Implementasi Data Mining menggunakan Weka. Malang : UB Press.
- [6]. Rajaguru, Harikumar., dan Suni Kumar Prabhakar. 2017. KNN Classifier and K-Means Clustering For Robust Classification of Epilepsy From EGG Signals. Hamburg : Anchor Academic Publishing.
- [7]. Irwansyah, Edy., dan Muhammad Faisal. 2015. Advanced Clustering Teori dan Aplikasi. Sleman : Deepublish.
- [8]. Mitra led, 2018, Definisi Lampu ,Jenis-Jenis dan Fungsinya, updated 10 October 2018, dilihat 10 April 2021, <<https://www.mitra-led.com/news/10/Definisi-Lampu-Jenis-Jenis-dan-Fungsinya>>.
- [9]. Flinsetyadi, 2020, Persiapan Data Dalam Data Mining: Data Transformation, updated 16 Juli 2020, dilihat 3 Agustus 2021, <<https://flinsetyadi.com/persiapan-data-dalam-data-mining-data-transformation/>>.

## BIOGRAPHY

**Prayogo Putra Tjaya**, lahir di Tangerang pada 14 September 1999. Menyelesaikan kuliah Strata I (S1) pada tahun 2021 pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.

**Rino S.Kom., M.Kom**, Saat ini bekerja sebagai dosen Tetap dan Kaprodi pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.

**Hartana Wijaya S.Kom., M.Kom**, Saat ini bekerja sebagai dosen Tetap pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.