

ANALISIS DAN PENERAPAN METODE KLASTERISASI MENGGUNAKAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN ORDER PRODUK PADA PT. ARTHAUTAMA PLASINDO

A. Yudi Permana¹), Muhtadin²)

Program Studi Teknik Informatika Fakultas Teknik
Univeristas Pelita Bangsa
yudi@pelitabangsa.ac.id

Disetujui, 30 Desember 2019

Abstraksi

Pemasaran dalam dunia perusahaan pada akhirnya menciptakan produk kepada konsumen di pasar. Penciptaan serta penyediaan produk tersebut didasarkan pada kebutuhan dan keinginan *customer*. PT. Artha Utama Plasindo adalah perusahaan yang bergerak pada bidang manufaktur injection molding part (komponen berbahan plastik) dan menjadi vendor dari berbagai macam *customer* baik dalam bidang manufaktur otomotif maupun elektronik. Salah satu *customer* besar dari PT. Artha Utama Plasindo adalah PT. Yamaha Indonesia Motor Mfg. (YIMM), namun pada kenyataannya di lapangan masih sering terjadi kendala dalam proses produksi maupun persediaan barang untuk *customer* YIMM tersebut. Diantara permasalahannya adalah terjadinya *over capacity* pada gudang penyimpanan serta banyaknya produk yang *shortage* maupun lebih *quantity*. Pada penelitian ini, peneliti akan memanfaatkan data *customer order* pada PT. Artha Utama Plasindo dengan menggunakan algoritma K-Means dalam penerapan klasterisasi yang diharapkan dapat menentukan tingkat pesanan produk PT. Yamaha Motor Indonesia Mfg pada PT. Artha Utama Plasindo untuk memberikan alternatif solusi terhadap kondisi permasalahan yang terjadi. Dalam penelitian ini untuk kategori kluster dibuat menjadi 3 *cluster*, yaitu tinggi untuk produk *High Demand*, sedang untuk produk *Medium Demand*, dan rendah untuk produk *Low Demand*. Hasil klasterisasi dari 33 record data yang dijadikan data training, ada sebanyak 5 produk yang masuk kluster pertama (tinggi), kemudian 6 produk di kluster kedua (sedang) dan 22 produk masuk dalam kluster ketiga (rendah).

Kata kunci: Produksi, penyimpanan, *over capacity*, algoritma K-Means

Abstract

Marketing in the corporate world ultimately creates products for consumers in the market. The creation and supply of these products is based on the needs and desires of the customer. PT. Artha Utama Plasindo is a company engaged in manufacturing injection molding parts (plastic components) and is a vendor of various customers in both automotive and electronic manufacturing. One of the big customers of PT. Artha Utama Plasindo is PT. Yamaha Indonesia Motor Mfg. (YIMM), but in reality in the field there are still often obstacles in the production process and inventory of goods for the YIMM customers. Among the problems are the occurrence of over capacity in the storage warehouse and the number of products that have shortage and more quantity. In this study, the authors will utilize customer order data at PT. Artha Utama Plasindo using the K-Means algorithm in the application of clustering which is expected to determine the order level of PT. Yamaha Motor Indonesia Mfg at PT. Artha Utama Plasindo to provide alternative solutions to the conditions of the problems that occur. In this study the cluster category was made into 3 clusters, namely high for High Demand products, medium for Medium Demand products, and low for Low Demand products. The results of the clustering of 33 data records used as training data, there are 5 products included in the first cluster (high), then 6 products in the second cluster (medium) and 22 products included in the third cluster (low).

Keyword: Production, storage, *over capacity*, K-Means algorithm

1. Pendahuluan

Dalam era globalisasi saat ini, perkembangan teknologi merupakan salah satu aspek yang dapat dimanfaatkan untuk berbagai macam kemudahan guna mencapai tujuan individu maupun organisasi, termasuk diantaranya mengolah berbagai jenis data dan informasi. Tentunya pengolahan data dan informasi ini juga didukung oleh perkembangan perangkat lunak (*software*) yang semakin hari semakin memberikan kenyamanan manusia dalam menyelesaikan setiap persoalan yang dihadapi.

Salah satu *customer* besar dari PT. Artha Utama Plasindo adalah PT. Yamaha Indonesia Motor Mfg. (YIMM). Hal ini tentu menjadi tantangan bagi PT. Artha Utama Plasindo selaku vendor untuk memberikan pelayanan yang baik, mulai dari segi kualitas produk yang dibuat, sampai dengan memastikan produk tersedia dan dapat dikirimkan sesuai dengan jadwal dan kebutuhan dari PT. YIMM selaku *customer*. Namun pada kenyataannya di lapangan masih sering terjadi kendala dalam proses produksi maupun persediaan barang untuk *customer*

YIMM tersebut. Beberapa hal yang biasa terjadi adalah *shortage* (kurang) atau lebih *quantity* pada produk tertentu dikarenakan masih belum optimalnya proses penjadwalan produksi. Selain itu pengelolaan stok yang tidak akurat juga menimbulkan penumpukan barang (*over capacity*) pada gudang penyimpanan yang berimbas pada biaya simpan yang tinggi dan tidak ekonomis. Kondisi ini juga menimbulkan dampak lain seperti kurang efektifnya proses FIFO (*First In First Out*) pada gudang *finish good* akibat manajemen penyimpanan barang yang tidak maksimal.

Dalam penelitian ini permasalahan akan dibatasi pada hal – hal berikut Analisis data hanya memakai jenis data yang berhubungan dengan *input* dan *output* gudang penyimpanan, dilihat dari total produksi dan total pengiriman produk dari *customer* PT. Yamaha Indonesia Motor Mfg. Penggunaan metode Klasterisasi dalam penelitian ini diukur dengan algoritma K-Means pada data yang dijadikan data training. Aplikasi yang dibuat berdasar data *customer* yang dianalisis, yaitu PT. Yamaha Indonesia Motor Mfg.

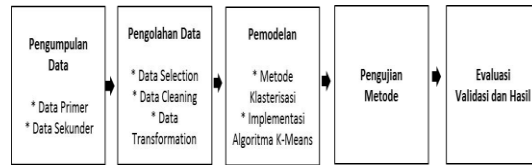
2. Tinjauan Studi

Penjelasan secara singkat teori-teori yang digunakan dalam melakukan penelitian yang terdiri dari Data mining, algoritma K-means, dan *domain knowledge*, serta kerangka berpikir dari penelitian yang dilakukan. (Beni MM & Herlina LS, 2015) menggunakan metode K-means dalam pengelompokan penjualan produk untuk mencari penjualan yang rendah dan penjualan yang tinggi pada swalayan FADHILA Bengkulu. Dari data yang diolah dengan sampel data yang diambil menghasilkan dua jenis kelompok tingkat penjualan, yaitu data penjualan rendah dan data penjualan tinggi. Sehingga dengan adanya pengelompokan data ini pihak swalayan Fadhilla dapat mengetahui jenis barang yang laris terjual dan tidak. (Gunawan A, Firman AP, Faiza R, 2016), metode K-Means dilakukan dengan beberapa tahapan untuk penerapan data mining dalam melihat klasifikasi potensi pemakaian air pelanggan di PDAM Tirta Raharja, dimana algoritma ini menunjukkan bahwa penerapan data mining pemakaian air pelanggan dengan algoritma ini memiliki penilaian fungsi-fungsi yang ada pada sistem ini yaitu 95,80% sistem dapat berjalan. (Muningsih E, Kiswati S, 2015).

Metode K-Means *Clustering* digunakan mengelompokkan produk yang dijual pada online shop Ragam Jogja menjadi beberapa *cluster* untuk mengetahui produk mana yang paling diminati, produk diminati, dan produk kurang diminati. Penelitian menggunakan metode ini karena dianggap paling populer dan salah satu metode terbaik dalam algoritma *clustering*. Hasil penelitian menunjukkan bahwa algoritma tersebut dapat menentukan kelompok produk yang masuk untuk jumlah stok banyak, sedang, dan sedikit berdasarkan transaksi penjualan. Dan dari 3 penelitian sebelumnya diatas maka dapat disimpulkan bahwa pemanfaatan metode Klasterisasi dengan menggunakan algoritma K-Means dapat membantu menyelesaikan masalah – masalah dalam menentukan kluster ataupun tingkat posisi tertentu dari sebuah data. Selain itu penerapan metode dan algoritma tersebut dapat digunakan pada berbagai bidang dengan tujuannya masing-masing. Maka pada penelitian ini akan digunakan kembali metode Klasterisasi dengan algoritma K-Means dengan penggunaan jenis dataset yang berbeda yaitu berupa data *customer order* yang meliputi jumlah produksi dan pengiriman perusahaan.

3. Desain Penelitian/ Metodologi

Dalam penelitian ini metode yang di gunakan yaitu metode data mining dengan laogaritma K-means dengan tahapan sebagai berikut. Tahap pengumpulan data pengolahan data tahap clustering maka dibuatlah langkah-langkah tersebut untuk meneliti dalam penelitian yaitu agar dapat mempermudah dan sesuai dengan tujuanyang diinginkan :



Gambar 1. Tahapan Penelitian

3.1 Pengumpulan data

Pada tahap ini dilakukan pengumpulan data yang dibutuhkan dalam penelitian. Jenis data yang digunakan adalah data primer dan data sekunder. Data Primer [a] Observasi adalah suatu cara untuk mendapatkan data dengan mengadakan dan pengamatan penelitian secara langsung di PT. AUP. Dalam penelitian ini, pendekatan observasi dilakukan pada bagian bagian terkait guna mengamati arus alur data yang digunakan. [b] Data dokumentasi ini diambil dari data *customer order* yang mencakup data produksi dan pengiriman pada PT.AUP. Data Sekunder [c] Tinjauan Pustaka Studi pustaka yang dilakukan memiliki acuan pada buku-buku serta jurnal-jurnal maupun literatur lainnya yang memiliki cakupan penelitian sejenis untuk mendukung teori yang digunakan serta menjadikan pembandingan dengan hasil yang didapat dari penelitian sebelumnya

3.2 Tahap Pengolahan Data

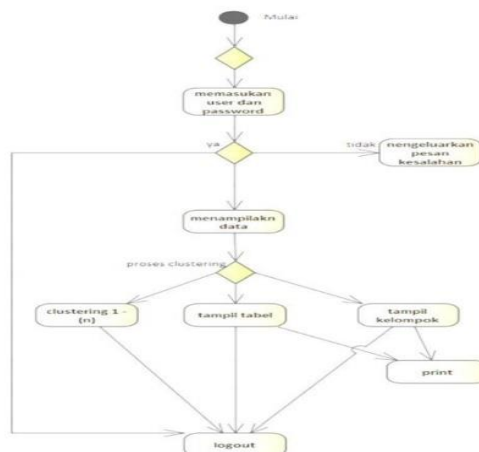
Menggambarkan data mining sebagai proses pencarian secara otomatis informasi yang berguna dalam tempat penyimpanan data berukuran besar. Istilah lain yang sering digunakan diantaranya *knowledge discovery in databases*, *knowledge extraction*, *data* atau *pattern analysis*, dan *business intelligence*.

Tahapan yang dilakukan dalam proses data mining untuk bidang industri menggunakan model CRISP-DM (Cross Industry Standard Process for Data Mining). CRISP-DM sangat sesuai digunakan untuk kebutuhan di industri karena tahapan yang dilakukan mudah dipahami dan sederhana. Pada tahap *data cleaning* merupakan proses pembersihan dari data yang akan dipakai untuk penghapusan data dengan membuang *missing value*, duplikasi data, dan memeriksa inkonsistensi data dan memperbaiki kesalahan pada data. Proses pembersihan data dilakukan secara manual dengan bantuan *software spreadsheet*. dan Ada 6 tahap pada model CRISP-DM (Alamsyah A & Rusdiawan WT, 2019), yaitu sebagai berikut : [a] *Business Understanding* [b] *Data Understanding* [c] *Data Preparation* [d] *Modelling* [e] *Evaluation* [f] *Deployment*

3.3 Tahapan Clustering

Clustering merupakan klasifikasi tanpa pengawasan adalah merupakan proses partisi sekumpulan objek dari data suatu set menjadi beberapa kelas.

Dalam penelitian ini peneliti menggunakan pendekatan Data Mining untuk melakukan *analisis data*. Data yang akan dijadikan dataset dalam penelitian ini adalah data produksi dan pengiriman produk yang didapatkan dalam bentuk file *spreadsheet* berformat excel sehingga tidak dapat langsung digunakan dikarenakan data masih terpisah pada beberapa kolom dan sheets



Gambar 2. Activity Diagram Aplikasi

4. Hasil Penelitian dan Pembahasan

Urutan langkah yang dilakukan dalam menggunakan algoritma K-Means adalah sebagai berikut Menentukan jumlah *cluster* yang akan dipakai, dalam penelitian ini dipakai 3 jenis *cluster*, antara lain High Demand (C1) untuk tingkat produksi tinggi, Medium Demand (C2) untuk tingkat produksi sedang, dan Low Demand (C3) untuk tingkat produksi rendah. Menentukan nilai centroid pada tahap awal untuk iterasi ke- 0, dilakukan secara random dengan rumus menentukan target awal k-means untuk mendapatkan target data atau jarak antar kelompok, menggunakan rumus berikut :

$$C_n \text{ Awal} = \frac{\text{Jumlah Data}}{\text{Jumlah Class} + 1}$$

Dengan menerapkan rumus diatas, maka didapatkan hasil sebagai berikut

$$C_n \text{ Awal} = \frac{33}{3+1} = 8.25$$

Maka ditetapkan untuk nilai awal centroid adalah 8,25

Menentukan titik awal dari setiap cluster. Pada penelitian ini, untuk titik awal cluster diambil secara random, yaitu dari nilai pada record ke -1, record ke 17, dan record ke 33. Berikut adalah hasil penentuan untuk titik awal untuk tahap iterasi ke -1 dalam bentuk tabel:

Tabel 1. Titik Pusat Awal

Titik Pusat Awal	Total Qty Produksi	Total Qty Pengiriman
Cluster ke-1	781.870	752.194
Cluster ke-2	54.664	54.628
Cluster ke-3	30	30

Menghitung jarak terdekat masing – masing record dengan nilai awal cluster, dengan menggunakan rumus *Euclidean Distance* seperti berikut :

$$D = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan :

- D* = *Euclidean Distance*
- I* = banyaknya objek
- (*x*,*y*) = koordinat objek
- (*s*,*t*) = koordinat centroid

Berikut adalah contoh untuk perhitungan *Euclidean Distance* untuk record pertama :

$$D(c1) = \sqrt{(781.780-781.780)^2 + (752.194-752.194)^2} = 0$$

$$D(c2) = \sqrt{(781.780-54.664)^2 + (752.194-54.628)^2} = 1.007.684$$

$$D(c3) = \sqrt{(781.780-30)^2 + (752.194-30)^2} = 1.084.908$$

Kelompokkan objek berdasarkan jarak ke *centroid* terdekat.

Jarak hasil perhitungan akan dilakukan perbandingan diantara 3 cluster dari data terdekat dengan *pusat cluster* dengan mengambil nilai terkecil, jarak ini menunjukkan bahwa data tersebut berada dalam 1 kelompok dengan pusat *cluster* terdekat. Secara umum berikut adalah matriks kelompok cluster yang di terjadi pada iterasi ke -1, dilambangkan dengan “T”.

Maka berikut perhitungan untuk menentukan nilai titik cluster baru dari iterasi ke – 1 :

$$\begin{aligned} & \text{C1 baru ke 1} \\ & (\text{Total Produksi}) = \frac{1.670.941}{3} = 556.980 \end{aligned}$$

$$\begin{aligned} & \text{C1 baru ke 1} \\ & (\text{Total Pengiriman}) = \frac{1.625.984}{3} = 541.995 \end{aligned}$$

$$\begin{aligned} & \text{C2 baru ke 1} \\ & (\text{Total Produksi}) = \frac{2.619.703}{16} = 163.731 \end{aligned}$$

$$\begin{aligned} & \text{C2 baru ke 1} \\ & (\text{Total Pengiriman}) = \frac{2.545.43}{16} = 159.089 \end{aligned}$$

$$\begin{aligned} & \text{C3 baru ke 1} \\ & (\text{Total Produksi}) = \frac{83.867}{14} = 5.991 \end{aligned}$$

$$\begin{aligned} & \text{C3 baru ke 1} \\ & (\text{Total Pengiriman}) = \frac{82.803}{14} = 5.915 \end{aligned}$$

Melalui beberapa tahapan, pada penjelasan di bagian pemodelan data, didapatkan hasil bahwa proses *clustering* dengan algoritma K-Means berhenti pada iterasi ke-4, karena posisi objek dari masing – masing *cluster* sudah tidak berubah dan mendapatkannilai yang optimal. Berikut adalah bentuk *cluster* yang didapat Cluster pertama memiliki pusat (495.967, 480.638), sehingga dapat diartikan pada *cluster* ini adalah kelompok produk yang memiliki tingkat order yang tinggi atau High Demand. Produk yang masuk dalam kategori ini adalah SWITCH CORD 3RS-00, VISOR B56, SWITCH CORD ASSY 3AY-00, COVER BTRY SUB ASSY, dan VISOR 2DS.

Cluster kedua memiliki pusat (227.101, 219.838), sehingga dapat diartikan pada rentang *cluster* ini adalah kelompok produk yang memiliki tingkat order yang sedang atau Medium Demand. Produk yang masuk dalam kategori ini adalah HOOK 1 N4, SWITCH CORD 437-01, VISOR B3F, SWITCH CORD 24.185, 24.175), sehingga dapat diartikan pada rentang *cluster* ini adalah kelompok produk yang memiliki tingkat order yang rendah atau Low Demand. Produk yang masuk dalam kategori ini adalah diantaranya SCREEN EROPA, SIDE COVER, AIR INTAKE, dan OIL LEVEL.

Tabel 2. Tabel Jumlah Objek Masing-Masing Cluster

Cluster	Jumlah Objek	Kelompok
Cluster 1 (495.967, 480.638)	5	High Demand
Cluster 2 (227.101, 219.863)	6	Medium Demand
Cluster 3 (24.185, 24.175)	22	Low Demand

4.1 Implementasi Sistem

Pada aplikasi ini akan menghasilkan pembentukan kelompok produk untuk tingkat status order yang dimiliki. Aplikasi dibuat untuk memudahkan dalam melihat dan memproses data dengan adanya bantuan antar muka pengguna (*Graphical User Interface*) sehingga penggunaannya dapat lebih optimal bagi *user*.

No.	Nama Produk	Nama Produk	Tahun Produk	Total Qty Persepsi	Qty 1	Qty 2	Qty 3	Kategori
1	SWITCHCORD	SWITCH CORD 3RS-00	05/05	750.154	368.013	191.905	1.625.984	H
2	VISOR	VISOR B56	05/05	160.134	38.151	160.643	397.248	H
3	COVER BTRY SUB ASSY	COVER BTRY SUB ASSY	05/05	415.267	415.265	188.774	360.289	M
4	SWITCH CORD ASSY	SWITCH CORD ASSY 3AY-00	05/05	418.203	355.105	129.263	279.473	M
5	HOOK 1 N4	HOOK 1 N4	05/05	380.078	198.047	180.275	102.205	M
6	SWITCH CORD 437-01	SWITCH CORD 437-01	05/05	330.172	213.015	116.619	448.204	M
7	VISOR B3F	VISOR B3F	05/05	230.245	200.073	148.597	360.202	M
8	SWITCH CORD 24.185	SWITCH CORD 24.185	05/05	215.550	122.941	20.946	200.705	M
9	SWITCH CORD 24.175	SWITCH CORD 24.175	05/05	221.112	172.018	27.558	200.960	M
10	OIL LEVEL	OIL LEVEL	05/05	194.020	498.962	45.006	190.708	M

Gambar 3. Tampilan Halaman Data



No.	Kode Produk	Nama Produk	Tingkat Order
1	3RS-H0306-00	BAND. SWITCH CORD (3RS-00)	High Demand
2	6S5-FE1AA-00	VISOR	High Demand
3	3AV-H8008-00	BAND SWITCH CORD ASSY (3AV-00)	High Demand
4	2PH-XK212-00	COVER, BTRY. SUB ASSY	High Demand
5	2ZS-FE1AA-00	VISOR	High Demand
6	2PH-FE257-00	HOOK, 1 (2PH)	Medium Demand
7	437-83836-01	BAND. SWITCH CORD (437-01)	Medium Demand
8	83F-FE1AA-00	VISOR	Medium Demand
9	3AV-H8006-10	BAND. SWITCH CORD (3AV-10)	Medium Demand
10	8N1-XK212-00	COVER BATTERY SUB ASSY	Medium Demand
11	8N0-FE381-00	WINDOW FRONT	Medium Demand
12	874-F807U-00	SCREEN ERRORA	Low Demand
13	437-83836-11	BAND. SWITCH CORD (437-11)	Low Demand

Gambar 4. Tampilan Menu

5. Kesimpulan

Metode klasterisasi dapat diterapkan dalam menganalisis data customer order yang ada di PT. Artha Utama Plasindo, dimana data yang diambil adalah data produksi dan data pengiriman untuk customer Yamaha. Algoritma K-Means yang digunakan menghasilkan sebuah wawasan baru, yaitu pengelompokan produk berdasarkan 3 cluster, cluster 1 merupakan produk High Demand yang terdiri dari 5 jenis produk, cluster 2 produk Medium Demand terdiri dari 6 jenis produk, dan cluster 3 produk Low Demand terdapat 22 jenis produk.

Aplikasi yang dibuat dapat digunakan untuk pengguna awam yang akan melakukan pemrosesan data dengan bantuan antar muka pengguna. Aplikasi ini membantu dalam pemrosesan data dari langkah awal hingga dapat menghasilkan suatu pengelompokan order produk dari datayang di proses.

Daftar Pustaka

- Suyanto (2017). Data Mining Untuk Klasifikasi Dan Klasterisasi Data, Bandung: Informatika.
- Abdillah, G, Putra A.F & Renaldi F (2016, Maret). Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di PDAM Tirta Raharja Menggunakan Algoritma K-Means, SENTIKA Yogyakarta, pp. 498 –
- Metisen, B.M., Sari, H.L. (2017, Februari). Analisis Clusteirng Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Penjualan Produk Pada Swalayan Fadhila, Jurnal Media Infotama, Vol.11 No.2, pp. 110 –118
- Dewi Eka Putri., S.Kom., M.Kom. (2015, Oktober). Metode Non Hierarchy Algoritma K-Means Dalam Mengelompokkan Tingkat Kelarisan Barang, Prosiding SENATKOM Padang, Vol.1, pp. 36- 41
- Muningsih E, Kiswati S (2015, Maret). Penerapan K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang, Jurnal Bianglala Informatika, Vol.3 No. 1, pp. 10 – 17
- Triyanto, A.W. (2015, April). Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk, Jurnal Simetris, Vol.6 No. 1, pp. 183 – 188
- Windarto, A.P (2017, November). Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering, Jurnal Techno.com, Vol.16 No. 4, pp. 348 - 357
- Rusdiawan, W.T., Alamsyah A (2019, April). Eksplorasi Data Pelanggan Untuk Kontekstual Marketing Voice Over Long Term Evolution PT. Telkomsel Menggunakan Metode Clustering K-Means, e-Proceeding of Management, Vol.6 No. 1, pp. 537 -544