



Analisis K-Means Clustering Pada Pengiriman Produk Bearing

Danendra Bima Adhi Pramana¹, Tifani Amalina², Riza Ibnu Adam³

Teknik Informatika, Universitas Singaperbangsa Karawang

Abstract

Received: 12 Agustus 2022
Revised: 16 Agustus 2022
Accepted: 20 Agustus 2022

Delivery of finished goods is carried out based on the incoming PO from the customer and received by the Marketing Department. The Marketing Department will make a DO based on the PO that has been received and then the DO will be forwarded to the Shipping Section to make a Packing List. The packing list that has been made will be validated by the section head and the items contained in the packing list will be prepared by the Prepare Subsection. The validated packing list will be printed and forwarded to the Delivery Subsection for delivery with the goods that are ready to be shipped. Shipping lines are made by the shipping coordinator based on the thoughts of the shipping coordinator, this causes the distance between one driver and another driver to be uneven. To overcome this problem, we need a method of dividing the task of delivering finished goods to customers. The method used is the K-Means Clustering Method. The K-Means Clustering method is a clustering method that partitions data into clusters so that data that have similarities are in the same cluster. In this study, the K-Means Clustering method was proven to be able to minimize the difference in distance between the driver and one PIC with the other driver and PIC. The difference in distance between the manual method and the K-Means Clustering method is 87,203 km.

Keywords: Information System, Finished Goods Delivery, Shipping, K-Means Clustering Method..

(*) Corresponding Author: danendra.bima18247@student.unsika.ac.id¹, tifani.amalina18091@student.unsika.ac.id², riza.adam@staff.unsika.ac.id³

How to Cite: Pramana, D. B., Amalina, T., & Adam, R. (2022). Analisis K-Means Clustering Pada Pengiriman Produk Bearing. *Jurnal Ilmiah Wahana Pendidikan*, 8(15), 128-137. <https://doi.org/10.5281/zenodo.7048988>

PENDAHULUAN

Metode *clustering* yang cukup dikenal adalah *K-Means Clustering*. *K-Means* merupakan metode pengklasteran yang memisahkan data ke dalam kelompok yang berbeda. *K-Means* mampu meminimalkan rata-rata jarak setiap data ke klasternya. (Agustin, Fitria, & S, 2015). *K-Means Clustering* merupakan salah satu metode pembagian data ke dalam *cluster-cluster* sehingga data yang memiliki kesamaan berada pada satu *cluster* yang sama dan data yang memiliki ketidaksamaan berada pada *cluster* yang lain (Rohmawati, Defiyanti, & Jajuli, 2015)

Bearing yang telah selesai diproduksi akan disimpan ke dalam gudang barang jadi yang nantinya akan dikirimkan ke konsumen sesuai dengan *Delivery Order*(DO) yang diterima oleh Bagian *Shipping*. *Delivery Order* dibuat berdasarkan *Purchase Order* (PO) yang diterima bagian Marketing. Bagian Marketing akan mengantarkan DO ke Bagian *Shipping* sebagai surat perintah pengiriman barang.

Pemilihan jalur pengiriman dibuat sesuai dengan posisi perusahaan yang ditujukan. Apabila dalam satu jalur pengiriman terdapat beberapa perusahaan yang dilewati maka dapat digabungkan dalam satu pengiriman. Waktu keberangkatan



dilakukan pada pukul 09.00 WIB sedangkan waktu kedatangan tidak bisa ditentukan karena kondisi di lapangan yang tidak dapat diprediksi.

Pengiriman dilakukan ke beberapa perusahaan yang tersebar di daerah Jabodetabek, Banten, dan Karawang sehingga tidak dapat jika hanya ditangani oleh satu sopir dan satu PIC saja, maka harus dilakukan pembagian jalur pengiriman yang dilakukan oleh sopir dan PIC sesuai dengan daerah yang dituju. Saat ini pembagian jalur pengiriman masih dilakukan secara manual, yaitu dibuatkan jadwal sesuai dengan pemikiran pimpinan sub bagian Delivery. Kelemahan dari pembagian secara manual di antaranya pembagian jarak yang tidak sama, beban sopir yang tidak merata, dan informasi jalur pengiriman masih ditulis dalam papan tulis dan laporan pengiriman dalam file tertulis.

Berdasarkan uraian di atas dilakukan penelitian pada jurnal ini dengan judul “Analisis K-Means Clustering Pada Pengiriman Produk Bearing”

METODE PENELITIAN

Metode K-Means Clustering

Dalam penelitian ini dijelaskan dengan rinci metode penelitian yang digunakan untuk perancangan dan implementasi metode *K-Means Clustering* pada pengiriman barang jadi pada Bagian *Shipping*. Tahapan proses pada penelitian ini dimulai dari data *input*, penentuan jumlah *cluster*, *inisialisasi centroid*, perhitungan jarak setiap data terhadap masing-masing *centroid*, mengelompokkan data berdasarkan jarak terdekat antara data dengan *centroid*, memperbaharui nilai *centroid*.

Tahapan dari metode *K-Means Clustering* adalah menentukan jumlah *cluster* yang akan dibentuk yaitu berdasarkan jumlah sopir yang ada, inisialisasi nilai *centroid*, menghitung jarak setiap data input terhadap masing-masing *centroid* menggunakan rumus jarak Euclidean (Euclidean Distance), mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil), memperbaharui nilai *centroid* dengan menghitung nilai rata-rata dari *cluster* yang bersangkutan, lalu melakukan pengulangan dari langkah ketiga-kelima hingga anggota tiap *cluster* tidak ada yang berubah



Figure 1. *K-Means Clustering*

HASIL DAN PEMBAHASAN

Metode K-Means Clustering

Metode *k-means* adalah algoritma yang mempartisi data ke dalam cluster-cluster sehingga data yang memiliki kesamaan berada pada satu cluster yang sama dan data yang memiliki ketidaksamaan berada pada cluster yang lain (Rohmawati, Defiyanti, & Jajuli, 2015):

- Menentukan *k* sebagai jumlah cluster yang ingin dibentuk.
- Membangkitkan nilai *random* untuk pusat *cluster* awal (*centroid*) sebanyak *k*
- Menghitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean* (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidean Distance*:

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

Keterangan:

x_i: data kriteria

μ_j: *centroid* pada cluster ke-*j*

- Memperbaharui nilai *centroid*. Menurut Rahman dkk (2017) nilai *centroid* baru di peroleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$C_k = \frac{1}{n_k} \sum d_i$$

Keterangan:

n_k = jumlah data dalam *cluster k*

d_i = jumlah dari nilai jarak yang masuk dalam masing-masing *cluster*.

- Melakukan perulangan dari langkah 2 hingga 5 sampai anggota tiap *cluster* tidak ada yang berubah.
- Hitung jarak setiap data ke masing-masing *centroid* menggunakan persamaan *Euclidean Distance* yaitu:

$$d(P, Q) = \sqrt{\sum_{j=1}^P (x_j(P) - x_j(Q))^2}$$

MSE (*Mean Squared Error*)

Mean Squared Error (MSE) adalah metode alternatif untuk mengevaluasi teknik peramalan masing-masing kesalahan (selisih data aktual terhadap data peramalan) dikuadratkan, kemudian dijumlahkan dan dibagi dengan jumlah data (Pramita & Tanuwijaya, 2010). MSE dapat dihitung dengan rumus:

$$MSE = \sum \left(\frac{y_t - x_t}{n} \right)^2$$

Keterangan:

y_t = permintaan pada periode t

x_t = ramalan untuk periode t

n = total jumlah periode.

Mean Absolute Percentge Error (MAPE)

Mean Absolute Percentge Error (MAPE) merupakan prosentase yang dihitung dari nilai absolut kesalahan di masing-masing periode dan dibagi dengan jumlah data aktual periode tersebut kemudian dicari rata-rata kesalahannya (Pramita & Tanuwijaya, 2010). MAPE dihitung dengan rumus:

$$MAPE = \sum \left| \frac{y_t - x_t}{\frac{y_t}{n}} \right|$$

Keterangan:

y_t = permintaan pada periode t

x_t = ramalan untuk periode t

n = total jumlah periode

|| = nilai absolut.

Mean Percentage Error (MPE)

Mean Percentage Error (MPE) dihitung dengan membagi kesalahan tiap periode dengan nilai aktual periode tersebut, kemudian di rata-ratakan. Jika pendekatan peramalan tidak bias, nilai yang dihasilkan akan mendekati nol (Pramita & Tanuwijaya, 2010). MPE dihitung dengan rumus:

$$MPE = \sum \frac{y_t - x_t}{\frac{y_t}{n}}$$

Keterangan:

y_t = permintaan pada periode t

x_t = ramalan untuk periode t

n = total jumlah periode.

Sum Square Error (SSE)

SSE adalah salah satu metode statistik yang dipergunakan untuk mengukur selisih total dari nilai sebenarnya terhadap nilai yang tercapai. Istilah SSE juga disebut sebagai *Summed Square of Residuals* (Hayatunnufus, Andrizal, & Yendri, 2016). Berikut adalah rumus perhitungan SSE.

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

Keterangan:

x = nilai aktual atau sebenarnya

y = nilai yang tercapai.

Cluster Metode K-Means Clustering

Sampel Data	Nama Perusahaan	x	y
1	PT Astra Honda Motor	-0,130467	0,175055
2	PT DanMotor Indonesia	-0,014235	-0,020844
3	PT Yamaha Indonesia Motor Mfg	-0,025660	-0,001771

1. Langkah pertama dalam pengklasteran yaitu, menentukan jumlah *k-cluster* yang dibentuk. *K-cluster* yang dibentuk yaitu sebanyak 3 *cluster* karena dalam sehari bisa mengirimkan sampai dengan 3 kali pengiriman.
2. Menentukan titik pusat *cluster*/menentukan nilai *centroid*. Inisialisasi K pusat *cluster* dilakukan dengan cara acak. Pusat-pusat *cluster* diberi nilai awal dengan angka-angka acak, dapat dilihat pada. Pada kasus ini yang terpilih sebagai titik *centroid* awal adalah sampel data 1,2, dan 3.

Sampel Data	Nama Perusahaan	x	y
1	PT Astra Honda Motor	-0,130467	0,175055
2	PT DanMotor Indonesia	-0,014235	-0,020844
3	PT Yamaha Indonesia Motor Mfg	-0,025660	-0,001771

Tabel I Penentuan Nilai *Centroid*

- Menghitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidean*. (Rohmawati, Defiyanti, & Jajuli, 2015)

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

Keterangan:

x_i : data kriteria.

μ_j : *centroid* pada *cluster* ke- j .

Contoh:

Diketahui sampel data 1 (-0,130467; 0,175055) =
 $\sqrt{(-0,130467 - (-0,130467))^2 + (0,175055 - 0,175055)^2} = 0$ (jarak dari sampel data 1 ke *centroid* 1)

Jarak dari sampel data 2 ke *centroid* 1 (-0,014235; -0,020844) \leftrightarrow (-0,130467; 0,175055)
 $= \sqrt{(-0,014235 - (-0,130467))^2 + (-0,020844 - 0,175055)^2} = 0,227786$

Jarak dari sampel data 3 ke *centroid* 1 (-0,025660; -0,001771) \leftrightarrow (-0,130467; 0,175055)
 $= \sqrt{(-0,025660 - (-0,130467))^2 + (-0,001771 - 0,175055)^2} = 0,205553$

Jarak dari sampel data 4 ke *centroid* 1 (-0,027013; 0,006607) \leftrightarrow (-0,130467; 0,175055)
 $= \sqrt{(-0,027013 - (-0,130467))^2 + (0,006607 - 0,175055)^2} = 0,19768$

Jarak dari sampel data 5 ke *centroid* 1 (-0,143046; 0,191265) \leftrightarrow (-0,130467; 0,175055)

$$\sqrt{(-0,143046 - (-0,130467))^2 + (0,191265 - 0,175055)^2} = 0,020518$$

Jarak dari sampel data 6 ke *centroid* 1 (-0,148694; 0,165645) \leftrightarrow (-0,130467; 0,175055)

$$\sqrt{(-0,148694 - (-0,130467))^2 + (0,165645 - 0,175055)^2} = 0,020513$$

Langkah selanjutnya yaitu pengelompokan berdasarkan jarak minimum terhadap *centroid*. Berdasarkan perhitungan di atas, dapat ditentukan data mana yang masuk ke dalam *cluster* 1, 2, dan 3

No	Nama Perusahaan	Lintang	Bujur	Jarak <i>Centroid</i> ke Titik Sampel			Kelompok Cluster
				1	2	3	
1	PT Astra Honda Motor	- 0,130467	0,175055	0	0,227786	0,205553	1
2	PT DanMotor Indonesia	- 0,014235	- 0,020844	0,307529	0	0,022233	2
3	PT Yamaha Indonesia Motor Mfg	- 0,025660	- 0,001771	0,202514	0,022233	0	3
4	PT Suzuki Indomobil Motor	- 0,027013	0,006607	0,19768	0,030279	0,008487	3
5	PT Yi Shen Industrial	- 0,143046	0,191265	0,020518	0,248158	0,225926	1
6	PT Asahi Denso Indonesia	- 0,148694	0,165645	0,020513	0,229907	0,207763	1

Tabel II Jarak Minimum Terhadap *Centroid*

- Menentukan nilai *centroid* yang baru dengan cara menghitung rata-rata dari *cluster* yang terkait dengan menggunakan rumus (Rahman, Wiranto, & Anggrainingsih, 2017):

$$C_k = \frac{1}{n_k} \sum d_i$$

Keterangan:

n_k = jumlah data dalam *cluster* k.

d_i = jumlah dari nilai jarak yang masuk dalam masing-masing *cluster*.

Contoh:

- Untuk mencari titik *centroid* yang baru, perhatikan keanggotaan dari tiap klaster
- *Cluster* 1 terdiri dari sampel data: ((-0,130467; 0,175055); (-0,143046; 0,191265); (-0,148694; 0,165645))

Maka titik *centroid* dari *cluster* 1 adalah:

$$x = \frac{(-0,130467 + (-0,143046) + (-0,148694))}{3} = -0,14073567$$

$$y = \frac{(0,175055 + 0,191265 + 0,165645)}{3} = 0,17732167$$

- *Cluster* 2 terdiri dari sampel data: ((-0,014235; -0,02084). Maka titik *centroid* dari *cluster* 2 adalah:

$$x = \frac{(-0,014235)}{1} = -0,014235$$

$$y = \frac{(-0,02084)}{1} = -0,02084$$

- *Cluster* 3 terdiri dari sampel data: ((-0,025660; -0,001771); (-0,027013; 0,006607))

Maka titik *centroid* dari *cluster* 3 adalah

$$x = \frac{(-0,025660 + (-0,027013))}{2} = -0,02634$$

$$y = \frac{(-0,001771 + 0,006607)}{2} = 0,002418$$

Berdasarkan perhitungan di atas maka didapatkan titik *centroid* yang baru yaitu pada

<i>Centroid</i>	x	y
1	-0,14073567	0,17732167
2	-0,014235	-0,02084
3	-0,02634	0,002418

Tabel III *Centroid* Baru

5. Lakukan perulangan dari langkah 3-5 hingga anggota tiap *cluster* tidak ada yang berubah.

Pada kasus ini perhitungan dilakukan sampai iterasi II, karena setelah iterasi II tidak ada perpindahan *cluster* sehingga ini merupakan hasil optimal yang

diperoleh dari metode *K-Means Clustering*, dan pengklasteran yang dihasilkan yaitu sesuai dengan

No	Nama Perusahaan	Lintang	Bujur	Jarak Centroid ke Titik Sampel			Kelompok Cluster
				1	2	3	
1	PT Astra Honda Motor	- 0,130467	0,175055	0,0105	0,2127	0,20161	1
2	PT DanMotor Indonesia	- 0,014235	- 0,020844	0,2351	0	0,02622	2
3	PT Yamaha Indonesia Motor Mfg	- 0,025660	- 0,001771	0,2129	0,02223	0,00424	3
4	PT Suzuki Indomobil Motor	- 0,027013	0,006607	0,2051	0,03028	0,00424	3
5	PT Yi Shen Industrial	- 0,143046	0,191265	0,0141	0,2481	0,22200	1
6	PT Asahi Denso Indonesia	- 0,148694	0,165645	0,0141	0,2299	0,20400	1

Tabel IV. Perhitungan Iterasi II

mengambarkan bahwa terdapat pengelompokan *cluster* yaitu *cluster* I yang memiliki anggota sampel data 1, 5 dan 6, *cluster* II yang memiliki anggota sampel data 2, *cluster* III yang memiliki anggota sampel data 3 dan 4.

KESIMPULAN

Berdasarkan hasil analisis data yang telah dilakukan dalam penelitian jurnal tersebut dapat disimpulkan beberapa hal yaitu: pengiriman barang jadi berdasarkan permasalahan yang ditemukan mampu meminimalisir perbedaan jarak antar sopir dan PIC yang satu dengan sopir dan PIC yang lainnya dengan mengimplementasikan metode *K-Means Clustering* yang terbukti dapat meminimalisir perbedaan jarak antar sopir dan PIC satu dengan sopir dan PIC yang lain. Selisih jarak antara metode manual dengan metode *K-Means Clustering* sebesar 87,203 km.

REFERENCES

- Haviluddin, S. J., Putra, G. M., Puspitasari, N., & Pakpahan, H. S. (2021). Implementasi Metode K-Means Untuk Pengelompokan Rekomendasi Tugas Akhir. *Jurnal Ilmiah Ilmu Komputer*, Vol 16, No1.
- Hayatunnufus, A., Andrizal, & Yendri, D. (2016, Maret 23). *Pendeteksi dan Verifikasi Tanda Tangan Menggunakan Metode Image Domain Spasial*.

Retrieved Maret 14, 2019, from Document Repository Universitas Andalas:
<http://repo.unand.ac.id/772/>

- Liliana, E., Rolly, I., & Andreas, H. (2020). Aplikasi Pemilihan Rute Pengiriman Barang pada Perusahaan Elektronik di Surabaya dengan Menggunakan Metode K-Means Clustering Dan Google Maps AP. 1-5.
- Prafiti, A. (2020). Implementasi Metode K-Means Clustering Pada Aplikasi Pengiriman dan Pemantauan Barang Kargo Di PT Angkasa Pura.
- Rahman, A. T., Wiranto, & Anggrainingsih, R. (2017). Coal Trade Data Clustering Using K-Means (Case Study PT Global Bangkit Utama. *Jurnal Ilmiah Teknologi dan Informasi*, 24-30.
- Rena, N., & Tobing, F. A. (2020). Analisis Cluster Dengan Menggunakan K-Means Untuk Pengelompokan Online Customer Review Pada Online Marketplace. *METHODIKA*, Vol.6, No1-5.