

KLASTERISASI PENJUALAN UNTUK MENENTUKAN BAHAN BANGUNAN TERLARIS DENGAN MENGGUNAKAN METODE *K-MEANS* DI UD MAJU BERSAMA

Desy Julika Sari¹, Wiwin Handoko^{2*}, Parini²

¹Mahasiswa Prodi Sistem Informasi, STMIK Royal

²Dosen Prodi Sistem Informasi, STMIK Royal

**email*: win.van.handoko@gmail.com

Abstract: UD Maju Bersama is a business engaged in sales. This business sells building materials and building material tools such as cement, zinc, iron, paint and so on. In this case, UD Maju Bersama in its business does not yet use a computer to run its system, therefore sales data and purchase data are not well structured, so that UD Maju Bersama has difficulty in determining which building materials are in demand. The purpose of this research is to make it easier for UD Maju Bersama to find out which building materials are in demand so that they can be restocked. Therefore, a system was created with the help of the K-Means Clustering method, this method is used because this method has the ability to group data with a large enough quantity of data in an efficient manner so that later it can find out which building materials are in demand. In this study, perform calculations by grouping the data into 3 clusters, namely sales that are not selling well, selling well and selling very well. So that the results of this study are data on 2 items of building materials that will be re-stocked, namely Cement 3 wheels and Semen Andalas. This system is intended for UD Maju Bersama in determining the building materials to be stocked or purchased in advance so that later there will be no accumulation of goods in the warehouse.

Keywords: Data Mining; K-Means Clustering; Sale.

Abstrak: UD Maju Bersama merupakan usaha yang bergerak dibidang penjualan. Usaha ini melakukan penjualan bahan-bahan bangunan dan alat-alat bahan bangunan seperti Semen, Seng, Besi, Cat dan lain sebagainya. Dalam hal ini UD Maju Bersama dalam bisnisnya belum menggunakan komputer untuk menjalankan sistemnya oleh karena itu data penjualan, data pembelian tidak tersusun dengan baik, sehingga UD Maju Bersama kesulitan dalam menentukan bahan bangunan yang sangat laris. Tujuan dari peneitian ini dilakukan agar dapat mempermudah UD Maju Bersama dalam mengetahui bahan bangunan yang sangat laris sehingga dapat distok ulang. Maka dari itu dibuat sistem dengan bantuan metode *K-Means Clustering*, metode ini digunakan karena metode ini memiliki kemampuan untuk mengelompokkan data dengan kuantitas data yang cukup besar dengan cara efisien sehingga nantinya dapat mengetahui bahan bangunan yang sangat laris. Dalam penelitian ini melakukan perhitungan dengan mengelompokkan data menjadi 3 kluster yaitu penjualan kurang laris, cukup laris dan sangat laris. Sehingga hasil dari penelitian ini yaitu didapat data 2 item bahan bangunan yang akan di stok ulang yaitu Semen 3 roda dan Semen Andalas. Sistem ini ditujukan agar UD Maju Bersama dalam menentukan bahan bangunan yang akan distok atau dibeli terlebih dahulu sehingga nantinya tidak akan terjadi penumpukan barang digudang.

Kata Kunci: Data Mining; K-Means Clustering; Penjualan.

PENDAHULUAN

Kemajuan teknologi informasi yang berkembang begitu pesat disegala bidang, seperti bidang ekonomi, industri, kesehatan dan sebagainya. Dengan adanya teknologi ini memberikan kemudahan dalam mencari informasi yang ingin didapatkan. dalam dunia bisnis kita dituntut untuk dapat melakukan inovasi terhadap bisnis yang dijalankan, untuk dapat meembangkan dunia bisnis diperlukan strategi dalam dunia bisnis seperti meningkatkan kuliatas produk, penambahan jenis produk dan memanfaatkan sosial media. Namun pemanfaatan teknologi ini belum dilakukan beberapa toko salah satunya toko di aek loba yaitu UD Maju Bersama.

UD Maju Bersama merupakan usaha yang bergerak dibidang penjualan. Usaha ini beralamat dijalan Lintas Sumatera, Aek Loba Pekan, Kec. Aek Kuasan, Kab. Asahan. Usaha ini melakukan penjualan bahan-bahan bangunan dan alat-alat bahan bangunan seperti Semen, Seng, Besi, Cat dan lain sebagainya.

. Dalam hal ini UD Maju Bersama dalam bisnisnya belum menggunakan komputer untuk menjalankan sistemnya, sehingga data-data penjualan, data pembelian tidak tersusun dengan baik. Seharusnya data-data yang dapat dikelola dengan baik dapat dimanfaatkan untuk mengetahui bahan bangunan apa saja yang terlaris dan tidak terlaris sehingga bahan bangunan mana yang harus didahulukan untuk dibeli dalam memenuhi stok agar tidak terjadi kehabisan stok. Maka dari itu dengan memanfaatkan sistem data mining kita akan mengetahui barang yang kurang laris, cukup laris dan sangat laris.

Data mining adalah istilah yang digunakan untuk menggambarkan menemukan pengetahuan dalam *database*. *Data mining* adalah proses mengekstraksi dan mengidentifikasi informasi yang berguna dan pengetahuan yang relevan dari *database* besar menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin [1]. Ada beberapa metode yang tersedia untuk *data mining*. Salah satunya adalah *K-Means clustering*.

K-Means clustering adalah metode yang digunakan untuk mengelompokkan kelompok objek menjadi kelompok-kelompok (jumlah bilangan bulat positif) berdasarkan atribut atau karakteristik yang sama [2].

Tujuan dari Clustering agar dapat meminimalisir fungsi objektif yang diatur dalam proses clustering dimana secara umum digunakan untuk meminimalisasikan variasi di dalam suatu *cluster* dan optimasi variasi antar *cluster* [3].

Sistem pengelompokan terhadap pengolahan data-data yang digunakan akan memudahkan pihak perusahaan untuk mengetahui bahan bangunan yang kurang laris, cukup laris dan sangat laris, sehingga bahan bangunan yang tersedia di gudang tidak menumpuk maupun mengalami kekurangan.

METODE

Metode yang digunakan yaitu Metode K-Means. Algoritma *K-Means* merupakan cara yang umum dan sederhana untuk digunakan, dikerenakan metode ini mampu untuk mengelompokkan data denngan jumlah data yang cukup besar dengan waktu yang relatif cepat dan efisien [4].

Selain itu Algoritma *K-Means* merupakan algoritma clustering yang berulang-ulang, algoritma *K-Means* dimulai dengan pemilihan secara acak untuk menentukan nilai *K*, dimana *K* yang artinya banyaknya cluster atau kelompok yang akan dibuat. Kemudian menetapkan nilai *K* secara random, sehingga nilai tersebut sementara menjadi nilai centroid atau nilai titik pusat. Untuk dapat menghitung jarak setiap data yang ada terhadap centroid dengan menggunakan rumus Euclidean hingga dapat ditemukan jarak yang paling dekat dari setiap centroid yang telah ditentukan. Kemudian klasifikasikan data dengan berdasarkan kedekatannya dengan centroid. Lakukan langkah tersebut dengan hingga nilai centroid tidak berubah [5].

Pada dasarnya metode *K-Means Clustering* ini digunakan untuk mengelompokkan sejumlah objek dengan atribut yang sama atau karakteristik dengan data-datanya.

Langkah-langkah dalam metode *K-Means* adalah sebagai berikut :

1. Tentukan nilai *k* nya sebagai jumlah cluster yang akan dibentuk
2. Tentukan titik pusat awal dari setiap cluster
3. Hitung jarak setiap data diinput masing-masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) samapi ditemukan jarak yang terdekat dari setiap data dengan centroid.

Berikut adalah persamaan Euclidean (Euclidean Distance) pada rumus (1):

$$D(x,y) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2} \dots\dots\dots(1)$$

Keterangan :

D : Jarak

X : Data

Y : Centroid / Cluster

4. Mengklasifikasi data berdasarkan kedekatannya dengan centroid.
5. Hitunglah kembali pusat kluster dengan anggota cluster yang sekarang. Pusat cluster ialah nilai rata-rata dari semua data objek dalam cluster tertentu.

Hitung lagi setiap objek memakai pusat kluster yang baru. Jika pusat cluster tidak berubah lagi maka proses klustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat kluster tidak berubah lagi [6].

METODOLOGI PENELITIAN

Dalam metodologi penelitian ada beberapa urutan kerja yang harus diikuti, urutan kerja ini merupakan langkah-langkah yang harus dikerjakan sesuai dengan pokok permasalahan yang dibahas dalam penelitian [7].

Adapun Langkah langkah yang harus dilakukan mulai dari masalah yang ditemukan sampai solusi dari permasalahan tersebut sehingga hasilnya dapat diterapkan dengan baik. Adapun kerangka kerja penelitian ini dapat dilihat pada gambar 1 yaitu sebagai berikut :



Gambar 1. Kerangka Kerja Pemikiran

Berdasarkan kerangka kerja pemikiran diatas maka dapat diuraikan sebagai berikut :

Tahap 1 Mempelajari Studi Literatur

Dalam tahap ini, peneliti mempelajari literatur yang terkait dengan data mining, clustering dan algoritma K-Means.

Tahap 2 Mengumpulkan Data

Pada tahap ini untuk dapat mengumpulkan data dilakukan dengan observasi langsung ke UD Maju Bersama. Data yang dikumpulkan yaitu data-data penjualan bahan bangunan di UD Maju Bersama, setelah data dikumpulkan maka melakukan pengelompokkan data sehingga dapat memutuskan masalah-masalah yang akan dipecahkan serta mencari solusinya.

Tahap 3 Merumuskan Masalah

Setelah melakukan pengumpulan data maka tahap selanjutnya yaitu untuk dapat merumuskan masalah berdasarkan permasalahan yang ada. Bagaimana keputusan yang akan diambil untuk mendapatkan output apa saja yang akan dihasilkan dari informasi yang akan dibangun dengan menggunakan metode K-Means.

Tahap 4 Menganalisa Dan Merancang Secara umum.

Pada tahap ini menjelaskan metode analisa yang dipakai dalam merancang secara umum apa saja yang akan dibeli oleh konsumen.

Tahap 5 Mentransformasikan Data

Pada tahap ini melakukan transformasi data dengan melakukan penginisialan data agar dapat memudahkan dalam melakukan pengujian.

Tahap 6 Melakukan Pengujian dengan menggunakan Metode K-Means

Setelah melakukan transformasi data maka tahap selanjutnya melakukan pengujian dengan menggunakan metode K-Means dimana menggunakan data yang didapatkan sehingga hasil yang sudah diuji akan dapat digunakan untuk mengambil kebijakan terhadap penjualan bahan bangunan.

Tahap 7 Menarik Kesimpulan

Pada tahap penarikan kesimpulan dilakukan bertujuan untuk dapat membandingkan hasil yang diperoleh dari tahap implementasi sistem yang dibuat secara manual dengan menggunakan python.

HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan Data Penjualan Bahan Bangunan pada Tahun 2021. Dalam melakukan perhitungan untuk dapat mengetahui bahan bangunan Kurang Laris, Cukup Laris dan Sangat Laris dengan menggunakan metode *K-Means Clustering* yang menggunakan Aplikasi Python agar dapat membantu pihak UD Maju Bersama untuk mengetahui bahan bangunan apa saja yang Kurang laris, Cukup Laris dan Sangat Laris yang nantinya akan di stok atau dibeli terlebih dahulu agar tidak terjadinya kekurangan maupun penumpukan persediaan bahan bangunan.

Dalam hal ini data yang digunakan untuk contoh dalam penelitian ini yaitu 10 data saja yang akan ditampilkan dalam penelitian ini. Berikut ini data yang digunakan terdapat pada tabel 1 antara lain yaitu:

Tabel 1. Data Bahan Bangunan UD Maju Bersama

No	Nama Barang	Stok	Jumlah Terjual
1	Semen 3 Roda	700	651
2	Semen Andalas	700	550
3	Semen Merah Putih	100	52
4	Semen Dynamik	100	46
5	Semen Holcim	250	150
6	Semen Gresik	200	100
7	Semen Padang	150	100
8	Semen Garuda	230	150
9	Seng 6x9	100	40
10	Seng 7x9	200	186

Dari data tersebut kemudian ditentukan 3 cluster yang digunakan sebagai acuan dalam iterasi awal pada metode *k-means*. Cluster ini digunakan untuk dapat mewakili 3 kluster yaitu Kurang Laris, Cukup Laris dan Sangat Laris sehingga dapat diamsumsikan

bahwa jumlah inialisasi centroid $K=3$. Algoritma pengelompokan *K-means* dapat ditingkatkan secara signifikan dengan memakai teknik inialisasi yang lebih baik, dan dengan mengulang (memulai kembali) algoritma tersebut [8].

Berikut ini analisa data dengan menggunakan Phyton untuk dapat mengetahui Bahan Bangunan Kurang Laris, Cukup Laris dan Sangat Laris dengan menggunakan Metode K-Means Clustering antara lain sebagai berikut:

Langkah pertama yaitu dengan meload data atau menampilkan data di aplikasi phyton dengan menggunakan text editor *Jupyter Notebook* seperti gambar berikut:

```
In [16]: #menampilkan dataset
import pandas as pd
import numpy as np

df = pd.read_csv('bahan bangunan.csv', delimiter=',')
df.head()
```

Out[16]:

	No	Nama barang	Stok	Jumlah Terjual
0	1	Semen 3 Roda	700	651
1	2	Semen Andalas	700	550
2	3	Semen Merah Putih	100	52
3	4	Semen Dynamik	100	46
4	5	Semen Holcim	250	150

Gambar 1. Meload Data

Langkah Kedua yaitu untuk dapat mengetahui deskripsi datanya:

```
In [2]: #menampilkan banyak data
df.describe()
```

Out[2]:

	No	Stok	Jumlah Terjual
count	70.000000	70.000000	70.000000
mean	35.500000	121.300000	67.242857
std	20.351085	123.544435	107.532326
min	1.000000	11.000000	1.000000
25%	18.250000	42.500000	11.250000
50%	35.500000	100.000000	30.000000
75%	52.750000	150.000000	69.000000
max	70.000000	700.000000	651.000000

Gambar 2. Deskripsi Data

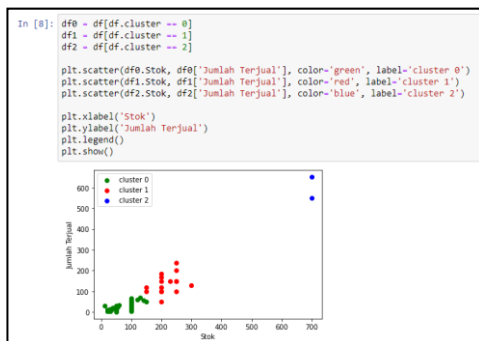
Langkah ketiga yaitu untuk mengetahui info data perihal type data yang digunakan:

```
In [3]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 70 entries, 0 to 69
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   No               70 non-null     int64
1   Nama barang     70 non-null     object
2   Stok             70 non-null     int64
3   Jumlah Terjual  70 non-null     int64
dtypes: int64(3), object(1)
memory usage: 2.3+ KB
```

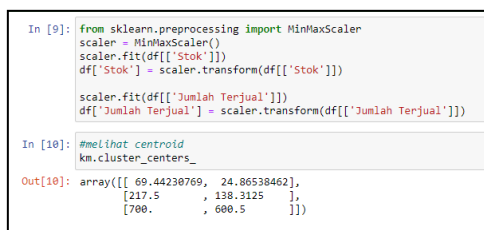
Gambar 3. Informasi Type Data

Langkah keempat yaitu menampilkan visualisasi datanya atau untuk dapat melihat persebaran data nya:



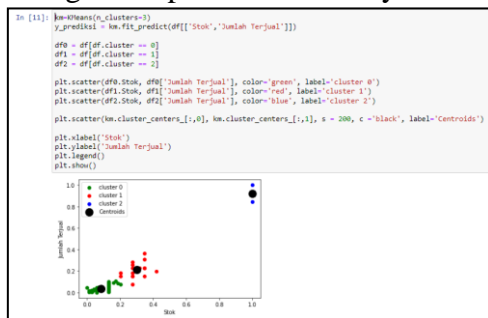
Gambar 8. Visualisasi Data

Langkah Kesembilan yaitu melihat nilai centroid atau nilai titik pusat yang digunakan sebagai nilai yang akan dijadikan sebagai kluster.



Gambar 9. Nilai Centroid

Langkah Kesepuluh yaitu menampilkan kembali visualisasi data hasil *cluster* dan *centroid* sehingga dapat mengetahui posisi *centroid* nya:



Gambar 10. Visualisasi Data Hasil Cluster dan Centroid

Langkah kesebelas yaitu melakukan group untuk dapat mengetahui nilai kluster disetiap kluster yang mana setiap kluster ituakan terlihat mana yang kluster yang Penjualan Kurang Laris, Penjualan Cukup Laris dan Penjualan Sangat laris.


```
In [12]: #melihat grup cluster agar dapat menentukan penjualan
#kurang Laris, cukup Laris dan sangat Laris
cluster_data = df[['Stok', 'Jumlah Terjual']]

kmeans_sel = KMeans(init = 'k-means++', n_clusters=3,
                    n_init=100, random_state=0).fit(cluster_data)
labels = pd.DataFrame(kmeans_sel.labels_)
clustered_data = cluster_data.assign(Cluster=labels)

In [13]: grouped_km = clustered_data.groupby(['Cluster']).mean().round(1)
grouped_km

Out[13]:
```

Cluster	Stok	Jumlah Terjual
0	0.1	0.0
1	0.3	0.2
2	1.0	0.9

Gambar 11. Group Cluster

Langkah Kedua belas yaitu mengganti nama cluster disetiap data, agar dapat lebih mudah mengetahui Penjualan Kurang Laris, Penjualan Cukup Laris dan Penjualan Sangat laris.

```
In [14]: #mengganti nama cluster
conditions = [
    (df['cluster'] == 0),
    (df['cluster'] == 1),
    (df['cluster'] == 2)
]
choices = ['Penjualan Kurang Laris', 'Penjualan Cukup laris', 'Penjualan Sangat Laris']
df['cluster'] = np.select(conditions, choices)
df

Out[14]:
```

No	Nama barang	Stok	Jumlah Terjual	cluster	
0	1	Semen 3 Roda	1.000000	1.000000	Penjualan Sangat Laris
1	2	Semen Andalas	1.000000	0.844615	Penjualan Sangat Laris
2	3	Semen Merah Putih	0.129173	0.076482	Penjualan Kurang Laris
3	4	Semen Dynamik	0.129173	0.069231	Penjualan Kurang Laris
4	5	Semen Hokim	0.346880	0.229231	Penjualan Cukup laris
...
65	66	Paku 2 Inchi	0.274311	0.183077	Penjualan Cukup laris
66	67	Paku 3 Inchi	0.274311	0.075385	Penjualan Cukup laris
67	68	Paku 4 Inchi	0.274311	0.152308	Penjualan Cukup laris
68	69	Triplex	0.201742	0.183077	Penjualan Cukup laris
69	70	Galon Penampung Air	0.056604	0.044615	Penjualan Kurang Laris

70 rows x 5 columns

Gambar 12. Mengganti Nama Cluster

Setelah melakukan perhitungan dengan menggunakan Phyton maka didapat hasil dari perhitungan yaitu penjualan yang sangat terlaris terdapat pada 2 item bahan bangunan yaitu Semen 3 Roda dan semen Andalas untuk itu maka UD Maju Bersama akan mengutamakan bahan bangunan 2 bahan bangunan yang sudah diketahui yang sangat terlaris sehingga dapat distok atau dibeli terlebih dahulu.

SIMPULAN

Berdasarkan hasil yang diperoleh menggunakan metode *K-Means* clustering maka hasilnya mendapatkan 3 kategori penjualan bahan bangunan yaitu kurang laris, cukup laris dan sangat laris. Hal ini dibuktikan berdasarkan perhitungan jarak terdekat berdasarkan penentuan nilai *centroid* secara random. Dengan mengambil data pada stok awal dan jumlah yang terjual sehingga hasil yang didapat berada pada 2 item yang nantinya distok lebih dulu.

Saran untuk peneliti berikutnya mungkin dapat melakukan perbandingan dengan menggunakan metode data mining lainnya dan menggunakan lebih banyak data lagi agar hasil yang didapat jauh lebih akurat.

DAFTAR PUSTAKA

- [1] A. Aqirati, “Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Klasterisasi Pola Penjualan Pestisida (Studi Kasus Di Toko Raja Bina Tani),” *J. Perencanaan, Sains, Teknol. dan Komput.*, vol. 2, no. 1, pp. 61–86, 2019.
- [2] M. H. Fakhriza and K. Umam, “ANALISIS PRODUK TERLARIS MENGGUNAKAN METODE K-MEANS CLUSTERING PADA ‘PT.SUKANDA DJAYA,’” *J. Inform.*, pp. 8–15, 2021.
- [3] P. D. P. Silitonga and M. Ginting, “Klasterisasi Keranjang Belanja DENGAN MENGGUNAKAN K-MEANS CLUSTERING,” *J. Mantik Penusa*, vol. 2, no. 2, pp. 164–168, 2018.
- [4] M. H. Siregar, “Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan),” *J. Teknol. Dan Open Source*, vol. 1, no. 2, pp. 83–91, 2018, doi: 10.36378/jtos.v1i2.24.
- [5] Y. D. Darmi and A. Setiawan, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk,” *J. Media Infotama*, vol. 12, no. 2, pp. 148–157, 2016, doi: 10.37676/jmi.v12i2.418.
- [6] Normah, S. Nurajizah, and A. Salbinda, “Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten,” *J. Tek. Komput. AMIK BSI*, vol. 7, no. 1, pp. 158–163, 2021, doi: 10.31294/jtk.v4i2.
- [7] F. Hadi and Y. Diana, “Pengklusteran Penjualan Bahan Bangunan Menggunakan Algoritma K-Means,” *JOISIE (Journal Inf. Syst. Informatics Eng.)*, vol. 4, no. 1, p. 22, 2020, doi: 10.35145/joisie.v4i1.629.
- [8] D. G. Ramadhan, I. Prihatini, and F. Liantoni, “Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means,” *Ultim. J. Tek. Inform.*, vol. 13, no. 1, pp. 33–38, 2021, doi: 10.31937/ti.v13i1.1981.