

Penerapan Metode *K-Means* untuk Mengklasifikasikan Penjualan Produk Olahraga Pada Toko Wan Toys & Sport

Marta Riama Uli Aritonang¹, Mhd. Anugrah Pramana², Putri Anggraini Dwiyantri³

^{1 2 3} Mahasiswa Program Studi Sistem Informasi, Universitas Royal

¹ martariamauli3@gmail.com; ² anugrahpermana330@gmail.com; ³ putrianggraini875@gmail.com

* Email Koresponden

INFO ARTIKEL

ABSTRAK

Kemajuan teknologi mendukung transformasi digital dalam pengelolaan data penjualan. Toko Wan Toys & Sport menghadapi kesulitan memahami pola penjualan, seperti bulan dengan penjualan tertinggi dan produk terpopuler. Pada penelitian ini untuk mengolah data memakai metode *K-Means clustering* melalui pendekatan CRISP-DM untuk mengelompokkan produk olahraga berdasarkan tingkat penjualannya. Hasil analisis menunjukkan bahwa metode ini mampu membagi produk dalam 3 kriteria: tinggi, sedang, dan rendah, sehingga memberikan wawasan strategis untuk pengelolaan stok dan pemasaran. Produk dengan penjualan tinggi diprioritaskan untuk stok, sedangkan produk dengan penjualan rendah menjadi target promosi. Metode ini efektif mendukung efisiensi operasional dan pengambilan keputusan berbasis data di toko Wan Toys & Sport.

Kata Kunci: K-Means, Klasifikasi penjualan, CRISP-DM, Produk Olahraga.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2024 by Author. Published by PT Beranda Teknologi Academia

ABSTRACT

Technological advances support digital transformation in sales data management. Wan Toys & Sport stores face difficulty understanding sales patterns, such as the highest sales months and most popular products. This research uses the K-Means clustering method with the CRISP-DM approach to group sports products based on their sales level. The analysis results show that this method is able to divide products into three categories: high, medium and low, thus providing strategic insight for stock management and marketing. Products with high sales are prioritized for stock, while products with low sales are targeted for promotion. This method effectively supports operational efficiency and data-based decision making at Wan Toys & Sport stores.

Keywords: K-Means, Sales Classification, CRISP-DM, Sports Products.

PENDAHULUAN

Dengan semakin pesatnya kemajuan teknologi, maka dapat dikatakan bahwa dunia sedang memasuki era transformasi digital, dimana teknologi telah berkembang dari skala besar menjadi mikro yang belum pernah kita gunakan sebelumnya hingga menjadi seperti sekarang ini[1]. Mengingat adanya teknologi ini, mampu mengembangkan dan memajukan serta meningkatkan penjualan sehingga dapat bermanfaat bagi pelaku usaha[2]. Pada penjualan yang memiliki peningkatan yang tinggi membutuhkan informasi serta identifikasi produk yang memiliki potensial dan produk yang tidak atau kurang memiliki potensial dalam penjualan[3]. Penjualan adalah salah

satu aspek terpenting dalam sebuah bisnis, jika tidak dijalankan dengan baik, hal ini dapat berarti bahwa bisnis tersebut berpotensi memberikan dampak negatif terhadap laba dan menurunkan pendapatannya[4]. Sama hal nya dengan Toko Wan Toys & Sport di Tanjung Balai.

Toko Wan Toys & Sport merupakan sebuah toko atau usaha yang menjual segala jenis alat dan perlengkapan olahraga, Toko Wan Toys & Sport berdiri pada tahun 2011 hingga saat ini, beralamatkan di Jalan Teuku Umar No. 55 Tanjung Balai, Pemilik toko tersebut bernama Lilis Suryani. Toko ini memiliki penjualan yang sangat baik dan dikatakan penjualan yang tinggi. Namun, di toko ini memiliki permasalahan dalam pengelolaan produk yang terjual tidak begitu baik dikarenakan pemilik toko tidak dapat mengetahui dalam beberapa bulan terakhir penjualan pada bulan apa saja yang memiliki peningkatan yang tinggi dan produk apa saja yang paling banyak diminati pada pelanggan. Oleh karena itu, sangat penting untuk mengelola penjualan agar produk yang dijual terkelola dengan baik dan dapat mempermudah pemilik toko dalam pekerjaan nya.

Maka, dari permasalahan tersebut diperlukan solusi untuk memecahkan masalah yang mana nanti jika solusi sudah di dapatkan maka akan meringankan pekerjaan pemilik toko, untuk itu diperlukan penggunaan data mining pada masalah tersebut. Penggunaan data mining menjadi proses yang digunakan yang memungkinkan dalam pengambilan keputusan dengan cepat dalam sebuah data mentah menjadi sebuah pengetahuan yang baru[5]. *Data mining* atau penambang data adalah proses untuk mengekstraksi pengetahuan dari kumpulan data yang tidak dapat dipahami dengan mudah melalui cara manual[6]. Data mining memiliki beberapa metode yang sering digunakan dalam mengelola data, adapun metode nya seperti *naive bayes*, *regresi linear*, *decision tree*, *K-means* dan *neural network*[7]. Pada permasalahan tersebut adapun metode yang diambil yaitu *K-means clustering*. Adapun definisi dari *K-means clustering* yaitu suatu metode yang dimiliki oleh data mining, yang mana metode ini melakukan proses pemodelan tanpa supervisi serta merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi[8]. Penggunaan metode *K-Means* dapat mengelompokkan dan memberikan informasi atau pengetahuan baru Bagi Toko Wan Toys & Sport mengenai penjualan pada bulan apa yang memiliki kenaikan dan barang apa yang lebih banyak terjual agar pemilik toko dapat memperbaiki kesalahan penjualan di bulan yang tidak memiliki kenaikan dan pemilik toko dapat menyediakan stok barang yang paling laku terjual[9].

METODE

Pada penelitian ini menggunakan *Machine Learning* sebagai dasar keilmuan, *Machine Learning* adalah salah satu cabang ilmu dari kecerdasan buatan (*Artificial Intelligence*). *Machine Learning* merupakan keahlian komputer untuk belajar melalui data tanpa harus memperlihatkan secara nyata atau spesifik apa yang harus dilakukan. Komputer memiliki kemampuan untuk belajar melalui data dan bisa menghasilkan prediksi, dan klasifikasi, serta keputusan[10]. *Cross-Industry Standard Process for Data mining*, atau CRISP-DM merupakan tehnik yang sering digunakan oleh para akademis dengan menggunakan proses pemodelan data dimaksud, Tujuannya adalah mengidentifikasi data yang menarik dan mempunyai kualitas yang baik.[11]. Tehnik *Cross-Industry Standard Process for Data mining* atau CRISP-DM dijelaskan dibawah ini. Tahap-tahap dari *K-means Clustering* merupakan langkah pertama dalam menentukan jumlah claster yang diinginkan, yang selanjutnya dengan menggunakan nilai k. Setelah itu, langkah berikutnya yaitu melakukan inisialisasi pusat claster (*centroid*) sejumlah k. Pada inisialisasi ini sering dilakukan secara acak dengan mengambil titik-titik awal melalui data yang tersedia. Berikutnya, perhitungan jarak antara setiap data *input* dengan masing-masing *centroid* dilakukan dengan memakai rumus jarak *Euclidean*. Rumus ini mengukur jarak antara data dengan *centroid* nya dengan persamaan matematis yang dijelaskan. *Euclidian Distance* memiliki persamaan, berikut ini adalah rumusnya:

$$\sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \dots \dots \dots (1)$$

Ket : De = Euclidian Distance

I = banyaknya objek

(x,y) = koordinat objek

(s,t) = koordinat *centroid*

Setelah jarak dihitung, data kemudian diklasifikasikan ke dalam *cluster* yang mempunyai *centroid* terdekat (dengan jarak terkecil). Langkah terakhir yaitu memperbaharui nilai-nilai *centroid*. Nilai-nilai *centroid* baru dihitung dengan mengambil rata-rata dari seluruh data yang sudah termasuk kedalam cluster tersebut, sesuai dengan rumus yang disebutkan. Berikut ini adalah rumus penentuan nilai *centroid*:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \dots \dots \dots (2)$$

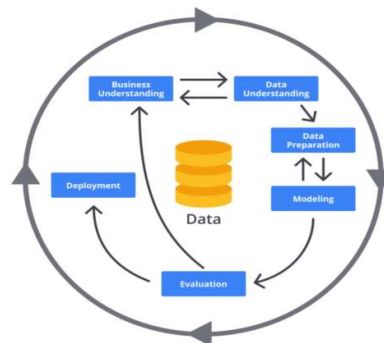
Ket : V_{ij} = *centroid* /rata-rata *cluster* ke-i untuk variabel ke-j

N_i = jumlah data yang menjadi anggota *cluster* ke-i

i,k = indeks dari *cluster*

j = indeks dari variabel

x_{kj} = nilai data ke-k yang ada di dalam *cluster* tersebut untuk variabel ke-j



Sumber [medium.easyread.co](https://medium.com/easyread.co)

Gambar 1. Tahapan CRISP-DM

Data Understanding

Data understanding merupakan langkah persiapan, melaksanakan pemeriksaan dari data yang dipakai, menyatukan data awal dan menjalankan untuk mengenali kualitas data. Data yang dipakai akan melalui tahap pemaparan dari semua fiturnya[12]. Pada data penelitian ini meliputi beberapa fitur yaitu nama barang, harga, jumlah transaksi, total penjualan, rata-rata. Berikut ini adalah understanding dataset yang digunakan :

```
[1]:
```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata
0	1	Piala Victori Set	135000	8	10	9.0
1	2	Piala Iner Set	130000	7	8	7.5
2	3	K.K Putsal Anti Slip	40000	10	14	12.0
3	4	Baju Putsal Set	55000	13	15	14.0
4	5	Bola Voli	77000	8	9	8.5
...
95	96	Sarung Samsak	80000	3	4	3.5
96	97	Protektor Badan Beladiri	100000	4	7	5.5
97	98	Pelindung Gigi	100000	2	2	2.0
98	99	Kasino Tape	30000	4	7	5.5
99	100	Papan Catur	60000	2	3	2.5

100 rows × 6 columns

Gambar 2. Deskripsi Data

Data Preparation

Data preparation adalah langkah persiapan data sebelum membuat analisis. Defenisi ini juga mencakup pemahaman akan pentingnya kesesuaian data dengan tujuan analisis, agar hasil analisis dapat dipercaya dan relevan. Dari defenisi yang jelas mengenai data preparation, akan mempermudah pelaksanaan proses persiapan data secara efektif dan efisien[13]. Berikut ini adalah proses melakukan data preparation :

```
[2]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype  
---  ---
0    No              100 non-null   int64  
1    nama_barang     100 non-null   object  
2    harga           100 non-null   int64  
3    jumlah_transaksi 100 non-null   int64  
4    total_penjualan 100 non-null   int64  
5    rata_rata       100 non-null   float64 
dtypes: float64(1), int64(4), object(1)
memory usage: 4.8+ KB
```

Gambar 3. Deskripsi Tipe Data

```
[3]: df.isnull().sum()

[3]: No              0
nama_barang         0
harga               0
jumlah_transaksi    0
total_penjualan     0
rata_rata           0
dtype: int64
```

Gambar 4. Cek Data Kosong (*Missing Value*)

Modelling

Modelling atau pemodelan merupakan tahapan implementasi algoritma yang akan digunakan untuk melakukan analisis, identifikasi, dan menghasilkan pola yang akan digunakan dalam data penelitian[14]. Tahapan pemodelan yang dilakukan menggunakan teknik klasifikasi dengan algoritma

K-means. *K-means clustering* adalah metode pengelompokan data menurut atribut yang sama dengan tujuan menemukan struktur tersembunyi dalam data tanpa label. Pengelompokan data bertujuan untuk meminimalkan fungsi objektif yang telah ditetapkan, yang pada dasarnya berupaya mengurangi variasi di dalam setiap kelompok.

Evaluation

Evaluation atau evaluasi merupakan suatu langkah dalam proses menganalisis hasil evaluasi suatu model yang telah diimplementasikan pada tahap pemodelan[15]. Temuan evaluasi ini menyoroti proses data mining yang telah selesai dan mengidentifikasi model yang paling cocok untuk digunakan.

Depeloment

Depeloment adalah proses pembuatan laporan atau artikel jurnal menggunakan hasil penelitian. Hasil dari penerapan metode *k-means* pada *data mining* di *interpretasikan* dalam bahasa pemrograman memudahkan Toko Wan Toys & Sport untuk membaca serta memahami informasinya dengan lebih cepat dan efektif.

HASIL DAN PEMBAHASAN

Tahap pertama adalah *import dataset* yang diperlukan untuk mengklasifikasikan produk olahraga.

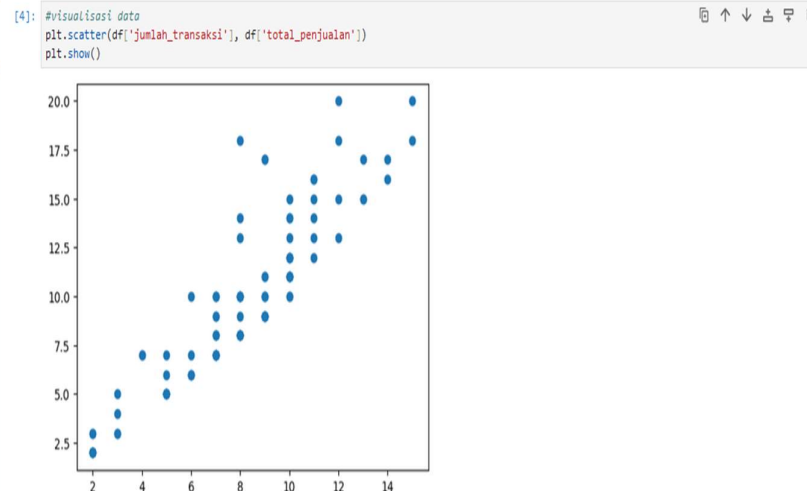
```
[1]: #Load library
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
#Load dataset
df = pd.read_excel('Data_Penjualan_WanToy.xlsx')
df
```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata
0	1	Piala Victori Set	135000	8	10	9.0
1	2	Piala Iner Set	130000	7	8	7.5
2	3	K.K Putsal Anti Slip	40000	10	14	12.0
3	4	Baju Putsal Set	55000	13	15	14.0
4	5	Bola Voli	77000	8	9	8.5
...
95	96	Sarung Samsak	80000	3	4	3.5
96	97	Protektor Badan Beladiri	100000	4	7	5.5
97	98	Pelindung Gigi	100000	2	2	2.0
98	99	Kasino Tape	30000	4	7	5.5
99	100	Papan Catur	60000	2	3	2.5

100 rows × 6 columns

Gambar 5. *Import Dataset* Produk Olahraga

Tahap kedua adalah visualisasi data pada tabel penjualan produk olahraga.



Gambar 6. Visualisasi *Dataset*

Tahap ketiga adalah membuat dataframe hasil *cluster*.

```
[7]: #membuat dataframe hasil kcluster agar bisa dilihat pada dataset
df['cluster'] = y_prediksi
df
```

```
[7]:
```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata	cluster
0	1	Piala Victori Set	135000	8	10	9.0	2
1	2	Piala Iner Set	130000	7	8	7.5	2
2	3	K.K Putsal Anti Slip	40000	10	14	12.0	0
3	4	Baju Putsal Set	55000	13	15	14.0	0
4	5	Bola Voli	77000	8	9	8.5	2
...
95	96	Sarung Samsak	80000	3	4	3.5	1
96	97	Protector Badan Beladiri	100000	4	7	5.5	1
97	98	Pelindung Gigi	100000	2	2	2.0	1
98	99	Kasino Tape	30000	4	7	5.5	1
99	100	Papan Catur	60000	2	3	2.5	1

100 rows × 7 columns

Gambar 7. *Dataframe Hasil Cluster*

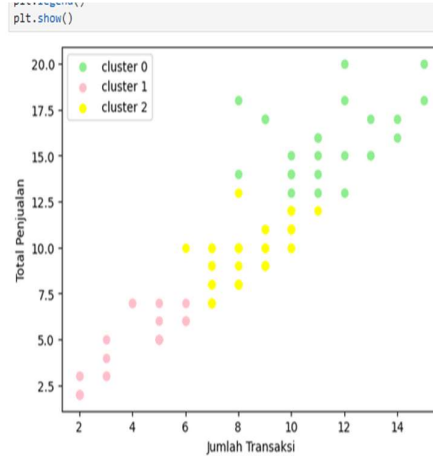
Tahap keempat adalah visualisasi hasil *cluster*.

```
[8]: #visualisasi hasil cluster

df0 = df[df.cluster == 0]
df1 = df[df.cluster == 1]
df2 = df[df.cluster == 2]

plt.scatter(df0.jumlah_transaksi, df0['total_penjualan'], color='lightgreen', label = 'cluster 0')
plt.scatter(df1.jumlah_transaksi, df1['total_penjualan'], color='pink', label = 'cluster 1')
plt.scatter(df2.jumlah_transaksi, df2['total_penjualan'], color='yellow', label = 'cluster 2')

plt.xlabel('Jumlah Transaksi')
plt.ylabel('Total Penjualan')
plt.legend()
plt.show()
```



Gambar 8. Visualisasi Hasil *Cluster*

Tahap kelima adalah *centroid* dari *dataset*.

```
[10]: #visualisasi hasil cluster

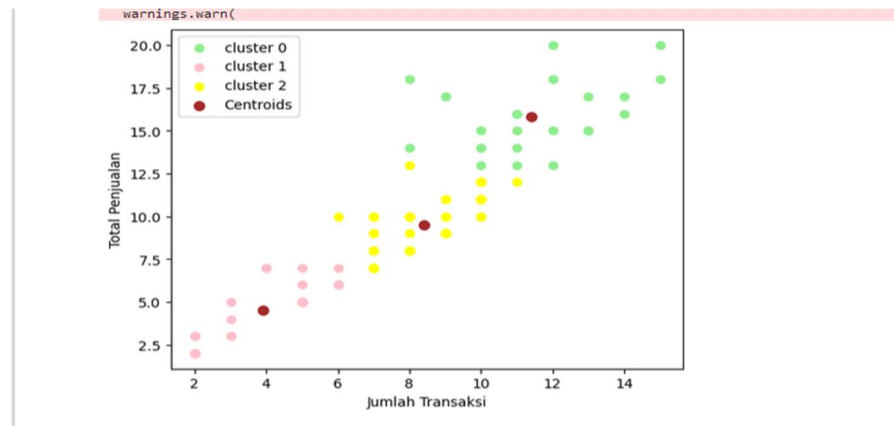
km = KMeans(n_clusters=3, random_state=0)
y_prediksi = km.fit_predict(df[['jumlah_transaksi', 'total_penjualan']])

df0 = df[df.cluster == 0]
df1 = df[df.cluster == 1]
df2 = df[df.cluster == 2]

plt.scatter(df0.jumlah_transaksi, df0['total_penjualan'], color='lightgreen', label = 'cluster 0')
plt.scatter(df1.jumlah_transaksi, df1['total_penjualan'], color='pink', label = 'cluster 1')
plt.scatter(df2.jumlah_transaksi, df2['total_penjualan'], color='yellow', label = 'cluster 2')

plt.scatter(km.cluster_centers_[0], km.cluster_centers_[1], s = 50, c = 'brown', label = 'Centroids')

plt.xlabel('Jumlah Transaksi')
plt.ylabel('Total Penjualan')
plt.legend()
plt.show()
```

Gambar 9. *Dataframe Hasil Cluster*

Tahap keenam adalah menampilkan *cluster* penjualan rendah.

```

•[32]: #Menampilkan Cluster Penjualan Rendah
df_rendah = df[df['cluster'] == 'Penjualan Rendah']
df_rendah

```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata	cluster
5	6	Bola Basket	200000	5	6	5.5	Penjualan Rendah
10	11	Ludo Ular Tangga	30000	6	6	6.0	Penjualan Rendah
11	12	Helm Bogo	135000	5	5	5.0	Penjualan Rendah
12	13	Dam Batu	60000	3	3	3.0	Penjualan Rendah
20	21	Gitar	300000	6	6	6.0	Penjualan Rendah
30	31	Pianika	120000	5	5	5.0	Penjualan Rendah
32	33	DP Tolak Peluru	350000	2	2	2.0	Penjualan Rendah
48	49	Tas Ransel	100000	6	6	6.0	Penjualan Rendah
50	51	Sepatu running Ortuseigt	250000	5	5	5.0	Penjualan Rendah
56	57	Rok Olahraga Wanita	50000	6	7	6.5	Penjualan Rendah
57	58	Sepatu Roda	220000	2	2	2.0	Penjualan Rendah
58	59	Papan Skateboard	150000	5	5	5.0	Penjualan Rendah
65	66	Sepatu Hiking	300000	5	5	5.0	Penjualan Rendah
67	68	Kajon	200000	5	5	5.0	Penjualan Rendah
84	85	Tenda Anak	180000	2	2	2.0	Penjualan Rendah

Gambar 10. *Cluster Penjualan Rendah*

Tahap ketujuh adalah menampilkan *cluster* penjualan sedang.


```
[28]: #Menampilkan Cluster Penjualan Sedang
df_sedang = df[df['cluster'] == 'Penjualan Sedang']
df_sedang
```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata	cluster
0	1	Piala Victori Set	135000	8	10	9.0	Penjualan Sedang
1	2	Piala Iner Set	130000	7	8	7.5	Penjualan Sedang
4	5	Bola Voli	77000	8	9	8.5	Penjualan Sedang
8	9	Baju Putsal Anak	50000	10	12	11.0	Penjualan Sedang
9	10	Piala Mangkok Set	150000	7	8	7.5	Penjualan Sedang
15	16	Lengging Nike	45000	11	12	11.5	Penjualan Sedang
16	17	Piala Victori Set	150000	9	9	9.0	Penjualan Sedang
17	18	Grip Raket Karet	15000	10	11	10.5	Penjualan Sedang
19	20	Sarung Tangan Kiper	220000	7	7	7.0	Penjualan Sedang
21	22	Kaca Mata Renang	30000	9	9	9.0	Penjualan Sedang
23	24	Raket Yonex CN	255000	7	7	7.0	Penjualan Sedang
24	25	Dekker	35000	8	10	9.0	Penjualan Sedang
25	26	Tali Gitar	50000	7	7	7.0	Penjualan Sedang
26	27	Bola Tennis Meja	30000	10	12	11.0	Penjualan Sedang

Gambar 11. Cluster Penjualan Sedang

Tahap kedelapan adalah menampilkan *cluster* penjualan tinggi.

```
[26]: #Menampilkan cluster Penjualan Tinggi
df_tinggi = df[df['cluster'] == 'Penjualan Tinggi']
df_tinggi
```

	No	nama_barang	harga	jumlah_transaksi	total_penjualan	rata_rata	cluster
2	3	K.K Putsal Anti Slip	40000	10	14	12.0	Penjualan Tinggi
3	4	Baju Putsal Set	55000	13	15	14.0	Penjualan Tinggi
6	7	Kaos kaki Panjang	30000	12	15	13.5	Penjualan Tinggi
7	8	Baju Bola Set	60000	15	18	16.5	Penjualan Tinggi
13	14	Bendera	50000	15	20	17.5	Penjualan Tinggi
14	15	Tas Bungkus Futsal	35000	14	16	15.0	Penjualan Tinggi
18	19	Grip Raket Handuk	20000	11	14	12.5	Penjualan Tinggi
22	23	Sepatu Futsal	150000	13	17	15.0	Penjualan Tinggi
28	29	Celana Training	50000	14	17	15.5	Penjualan Tinggi
31	32	Celana Pendek	35000	12	13	12.5	Penjualan Tinggi
33	34	Baju Club Anak	50000	10	15	12.5	Penjualan Tinggi
34	35	Baju Jersey GO	60000	9	17	13.0	Penjualan Tinggi
37	38	Sun Power	130000	11	16	13.5	Penjualan Tinggi
40	41	K.K Sambung	35000	12	18	15.0	Penjualan Tinggi

Gambar 12. Cluster Penjualan Tinggi

KESIMPULAN

Kemajuan teknologi yang semakin pesat telah mendorong perlunya transformasi digital, termasuk dalam pengelolaan data penjualan. Penelitian ini berhasil mengimplementasikan metode *K-means clustering* untuk mengelompokkan data penjualan di Toko Wan Toys & Sport berdasarkan pola tertentu, seperti bulan dengan penjualan tertinggi dan produk paling diminati. Dengan menggunakan pendekatan CRISP-DM, proses mulai dari pemahaman data hingga pengembangan aplikasi berbasis hasil penelitian telah dilakukan secara sistematis. Hasil penelitian menunjukkan bahwa metode ini dapat memberikan wawasan baru kepada pemilik toko, seperti identifikasi produk yang paling potensial dan waktu terbaik untuk meningkatkan strategi pemasaran. Prospek

pengembangan penelitian ini mencakup penerapan metode data mining lainnya, seperti *neural network* atau *decision tree*, untuk analisis yang lebih mendalam. Selain itu, aplikasi berbasis data mining ini memiliki potensi untuk diintegrasikan ke dalam sistem inventori toko sehingga mempermudah pengelolaan stok dan perencanaan penjualan di masa depan. Dengan implementasi yang lebih luas, metode ini dapat diterapkan di berbagai jenis usaha kecil hingga menengah, membantu mereka memanfaatkan teknologi untuk pengambilan keputusan yang lebih efektif dan efisien.

DAFTAR PUSTAKA

- [1] H. Abijono, P. Santoso, and N. L. Anggreini, "Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data," *J. Teknol. Terap. G-Tech*, vol. 4, no. 2, pp. 315–318, 2021, doi: 10.33379/gtech.v4i2.635.
- [2] A. R. Riefnaldi, A. Aranta, and M. Muaidi, "Pembuatan Sistem Informasi Pengarsipan Surat Pada Kantor Desa Sandik Berbasis Website," *J. Begawe Teknol. Inf.*, vol. 2, no. 2, pp. 191–202, 2021, doi: 10.29303/jbegati.v2i2.557.
- [3] M. R. Sulistio, N. Suarna, and O. Nurdian, "Analisa Penerapan Metode Clustering X-Means Dalam Pengelompokan Penjualan Barang," *J. Teknol. Ilmu Komput.*, vol. 1, no. 2, pp. 37–42, 2023, doi: 10.56854/jtik.v1i2.49.
- [4] N. P. Gantara and I. Ali, "Penerapan Metode K-Means Clustering Pada Penjualan Barang Di Sports Station," *E-Link J. Tek. Elektro dan Inform.*, vol. 18, no. 1, p. 28, 2023, doi: 10.30587/e-link.v18i1.5339.
- [5] S. W. Harjono, N. W. Utami, and I. G. A. P. D. Putri, "Klasterisasi Tingkat Penjualan pada Startup Panak.id dengan Algoritma K-Means," *J. Ilm. Teknol. Inf. Asia*, vol. 17, no. 1, p. 55, 2023, doi: 10.32815/jitika.v17i1.888.
- [6] K. Annisa, B. S. Ginting, and M. A. Syar, "Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada Pdam Langkat," *J. Sist. Inf. Kaputama*, vol. 6, no. 2, pp. 165–179, 2022, doi: 10.59697/jsik.v6i2.167.
- [7] M. A. P. P. Althaf and Rianto, "Analisis Probabilitas Bencana Alam dengan Penerapan Data Mining Menggunakan K-Means dan Linier Regression," *J. Mhs. Tek. Inform.*, vol. 2, no. 2, pp. 1–7, 2023, doi: 10.35473/jamastika.v2i2.2345.
- [8] R. A. Kurniawan, M. S. Hasibuan, P. Piramida, and R. S. Ramadhan, "Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara," *J. Comput. Sci. Informatics Eng.*, vol. 01, no. 1, pp. 10–18, 2022, doi: 10.55537/cosie.v1i1.27.
- [9] A. Andriani *et al.*, "Penerapan Algoritma K-Means Untuk Mengklasifikasikan Penjualan Produk Dettol," *J. Artif. Intell. Data Eng.*, vol. 1, no. 1, pp. 1–8, 2024.
- [10] M. A. R. Indra Budi Trisno, "WEBINAR ARTIFICIAL INTELLIGENCE DAN MACHINE LEARNING," *WEBINAR Artif. Intell. DAN Mach. Learn.*, vol. 2, no. 11, pp. 2307–2315, 2023, doi: 10.31862/9785426311961.
- [11] F. Febriansyah and S. Muntari, "Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 66–77, 2023, doi: 10.14421/jiska.2023.8.1.66-77.
- [12] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, "Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM," *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [13] R. A. Bakti, R. Wulanningrum, W. Cahyo, P. Jasmani, T. Informatika, and C. Author, "Sosialisasi dan Pelatihan Modul Data Preparation Aplikasi JIMAT di NPCI Kabupaten Kediri," vol. 4, no. 1, 2024.
- [14] A. Hardirega and I. Jaelani, "IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK (CNN) KLASIFIKASI MOTIF BATIK MENGGUNAKAN EFFICIENTNET-B1," vol. 8, no.



- 5, pp. 10023–10028, 2024.
- [15] Y. D. Pradvenanta and R. Prathivi, “Implementasi K-Means Untuk Pengelompokan Makanan Cepat Saji Bagi Penderita Penyakit Obesitas,” vol. 6, no. 1, 2024, doi: 10.47065/bits.v6i1.5279.