

PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK ANALISA PENJUALAN PADA TOKO NUSA INDAH

Erlinda¹, Helpi Nopriandi², Dwipa Junika Putra³

^{1,2}Program Studi Teknik Informatika Fakultas Teknik Universitas Islam Kuantan Singingi

³Program Studi Informatika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Syiah Kuala Banda Aceh

^{1,2} Jl. Gatot Subroto KM. 7 Kebun Nenas, Teluk Kuantan, Riau, telp. 0823 8746 6901

³ Jl. Teuku Nyak Arief No. 441, Kopelma Darussalam, Kec. Syiah Kuala, Kota Banda Aceh

Email: ¹erlinda120015@gmail.com, ²helpinopriandi83@gmail.com
³dwipajunikaputra@usk.ac.id

Abstrak

Toko Nusa Indah merupakan toko ritel yang menjual berbagai macam produk seperti sepatu, tas, dan aksesoris lainnya. Seiring bertambahnya volume transaksi, data penjualan yang dihasilkan semakin besar dan kompleks. Namun, data tersebut belum dimanfaatkan secara optimal untuk mendukung pengambilan keputusan bisnis. Penelitian ini bertujuan untuk menerapkan metode data mining dengan algoritma K-Means Clustering dalam menganalisis data penjualan guna mengelompokkan produk berdasarkan tingkat penjualan. Tahapan yang dilakukan meliputi pengumpulan data transaksi, preprocessing data, dan proses klasterisasi menggunakan algoritma K-Means. Hasil dari klasterisasi menunjukkan bahwa produk dapat dikategorikan ke dalam beberapa cluster, seperti produk dengan penjualan tinggi, sedang, dan rendah. Informasi ini sangat bermanfaat bagi manajemen toko dalam menyusun strategi pemasaran, perencanaan stok barang, serta evaluasi kinerja produk. Dengan demikian, penerapan metode K-Means Clustering dapat memberikan wawasan yang lebih mendalam dalam pengelolaan penjualan dan pengambilan keputusan di Toko Nusa Indah.

Kata Kunci: Data Mining, K-Means Clustering, Analisis Penjualan, Klaster Produk, Toko Nusa Indah

Abstract

Toko Nusa Indah is a retail store that sells a variety of products such as shoes, bags, and other accessories. As the volume of transactions increases, the resulting sales data becomes larger and more complex. However, this data has not been optimally utilized to support business decision-making. This study aims to apply data mining techniques using the K-Means Clustering algorithm to analyze sales data and group products based on their sales performance. The stages carried out include data collection, data preprocessing, and clustering using the K-Means algorithm. The clustering results show that products can be categorized into several clusters, such as high, medium, and low sales. This information is highly beneficial for store management in developing marketing strategies, inventory planning, and evaluating product performance. Thus, the application of the K-Means Clustering method can provide deeper insights into sales management and decision-making at Toko Nusa Indah.

Keywords: Data Mining, K-Means Clustering, Sales Analysis, Product Clustering, Toko Nusa Indah

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat saat ini telah mendorong pelaku usaha untuk memanfaatkan data sebagai dasar dalam pengambilan keputusan yang lebih tepat dan strategis. Salah satu teknologi yang dapat digunakan untuk menggali informasi berharga dari data adalah data mining. Data mining memungkinkan pengolahan data dalam jumlah besar untuk menemukan pola-pola tersembunyi yang sebelumnya tidak terlihat secara langsung, yang dapat dijadikan acuan dalam pengembangan strategi bisnis [1].

Toko Nusa Indah merupakan salah satu toko ritel yang berlokasi di Lubuk Jambi, Kecamatan Kuantan Mudik, Kabupaten Kuantan Singingi. Toko ini menjual berbagai jenis produk seperti tas, sepatu, dan

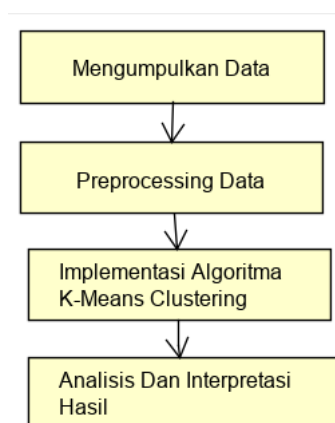
aksesoris lainnya. Seiring dengan meningkatnya jumlah transaksi setiap harinya, Toko Nusa Indah menghasilkan data penjualan yang cukup besar. Namun, data tersebut selama ini belum dimanfaatkan secara maksimal untuk mendukung pengambilan keputusan bisnis, seperti dalam penentuan stok barang, evaluasi performa produk, maupun strategi promosi.

Salah satu metode dalam data mining yang dapat digunakan untuk menganalisis dan mengelompokkan data adalah K-Means Clustering. Metode ini bekerja dengan cara membagi data ke dalam beberapa kelompok (cluster) berdasarkan kemiripan karakteristik tertentu. Dalam konteks penjualan, metode ini dapat membantu toko dalam mengelompokkan produk berdasarkan tingkat penjualan, sehingga dapat diketahui produk mana yang memiliki penjualan tinggi, sedang, maupun rendah [2].

Dengan penerapan metode K-Means Clustering, diharapkan Toko Nusa Indah dapat memperoleh wawasan yang lebih mendalam terkait pola penjualan produknya. Informasi ini akan sangat berguna dalam pengambilan keputusan strategis seperti perencanaan stok, penentuan produk unggulan, hingga pengembangan strategi pemasaran yang lebih efektif.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining yang menerapkan algoritma *K-Means Clustering* untuk menganalisis data penjualan pada Toko Nusa Indah. Tujuan dari penelitian ini adalah untuk mengelompokkan produk-produk yang dijual berdasarkan tingkat penjualannya agar manajemen toko dapat mengambil keputusan yang lebih tepat dalam hal pengelolaan stok dan strategi penjualan. Berikut Langkah-langkah Penelitian Penerapan Algoritma *K-Means Clustering*:



Gambar 1. Langkah-Langkah Penelitian

Keterangan Gambar 1. Langkah-langkah penelitian :

1. *Mengumpulkan Data*
Tahap pertama dalam proses data mining. Data dikumpulkan dari sumber yang relevan (misalnya, data penjualan dari Toko Nusa Indah) yang akan digunakan sebagai dasar untuk analisis selanjutnya.
2. *Preprocessing Data*
Proses pembersihan dan persiapan data sebelum dilakukan pemodelan. Meliputi: Menghapus data duplikat atau kosong, normalisasi data (jika diperlukan), Seleksi atribut yang relevan untuk clustering
3. *Implementasi Algoritma K-Means Clustering*
Di tahap ini, algoritma K-Means diterapkan untuk mengelompokkan data ke dalam sejumlah klaster berdasarkan kesamaan nilai antar atribut (misalnya stok awal, stok keluar, stok akhir).
4. *Analisis dan Interpretasi Hasil*
Tahap akhir berupa analisis hasil klaster yang terbentuk: Menilai karakteristik setiap klaster (misal: produk fast-moving dan slow-moving) Mengambil keputusan strategis dari hasil analisis (misalnya: strategi promosi atau pengelolaan stok)

2.1. Data Mining

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam *database*. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar [3].

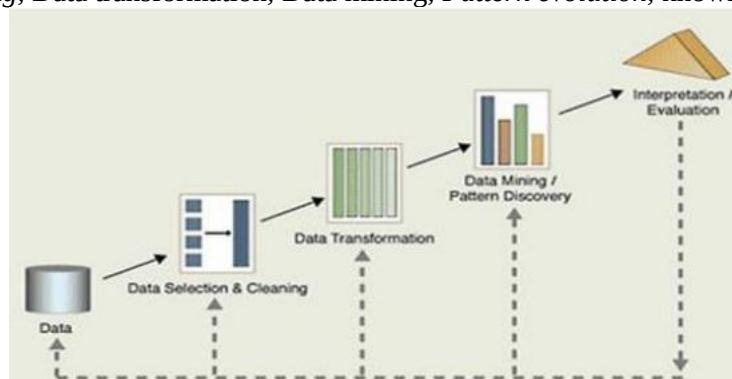
2.2. Algoritma K-Means

Algoritma *K-Means* adalah metode klasterisasi yang mengelompokkan data ke dalam *k cluster* berdasarkan jarak terdekat terhadap titik pusat (*centroid*). Langkah-langkah dasar dari algoritma ini meliputi:

1. Menentukan jumlah cluster (*k*).
2. Menginisialisasi *centroid* awal secara acak.
3. Mengelompokkan data berdasarkan jarak terdekat ke *centroid*.
4. Menghitung ulang posisi *centroid* dari masing-masing cluster.
5. Mengulangi proses hingga posisi *centroid* tidak lagi berubah signifikan atau jumlah iterasi maksimum tercapai.

2.3 Metode Penelitian

Metode pengembangan data mining yang digunakan untuk menganalisis data dalam penerapan data mining ini menggunakan proses tahapan *knowledge discovery in databases* (KDD) yang terdiri dari Data, Data *Cleaning*, Data transformation, Data mining, *Pattern evolution*, knowledge [4]:



Gambar 2. Proses KDD [4]

Berikut Keterangan gambar 2. Proses KDD (*knowledge discovery in databases*) diatas :

1. Data
Tahap awal berupa kumpulan data mentah dari berbagai sumber seperti *database*, data *warehouse*, *file log*, dan lain-lain. Yang bertujuan untuk menyediakan bahan dasar untuk analisis lebih lanjut.
2. Data selection dan *Cleaning*
Proses memilih *subset* data yang relevan dan membersihkan data dari *noise*, kesalahan, atau data yang tidak lengkap. Guna untuk meningkatkan kualitas data agar analisis lebih akurat dan bermakna.
3. Data Transformasi
Data yang telah dibersihkan kemudian ditransformasikan ke dalam format atau struktur yang sesuai untuk proses mining, seperti normalisasi, agregasi, atau encoding. Guna untuk menyesuaikan format data agar lebih siap untuk diproses oleh algoritma data mining.
4. Data Mining/ *Pattern Discovery*
Proses inti dari data mining di mana algoritma digunakan untuk menemukan pola, korelasi, atau tren dalam data. Metode umum: klasifikasi, klustering, asosiasi, regresi, dan lainnya. Bertujuan untuk menghasilkan informasi atau pola tersembunyi dari data yang telah dipersiapkan.
5. Interpretation/Evaluation

Tahap terakhir di mana hasil dari data mining dievaluasi dan ditafsirkan agar dapat digunakan untuk pengambilan keputusan. Bertujuan untuk memastikan bahwa pola yang ditemukan bermakna, relevan, dan berguna bagi bisnis atau penelitian.

3. HASIL DAN PEMBAHASAN

3.1 Data

Data penjualan tersebut ditarik menggunakan teknik data primer yang artinya data diambil langsung dari penjualan Toko Nusa Indah Lubuk Jambi, adapun data dilihat pada tabel berikut ini:

Tabel 1. Data Penjualan

No	Nama Barang	Kode Barang	Stok		
			Awal	Keluar	Akhir
1	Sandal ando hawai	01	45	35	10
2	sandal grania nininn NI	02	68	36	32
3	sandal youth balance kning qpr	03	90	47	43
4	sandal karet fimio mntira tbi smg	04	78	75	3
5	sandal youth balance manik cwek	05	100	80	20
6	ransel sekolah import chibao	06	89	78	11
7	ransel santai import	07	88	79	9
8	tas santai lokal	08	99	67	32
9	tas import hermes	09	78	55	23
10	dompet meet you forever	10	70	48	22
11	dompet kartu	11	89	76	13
12	dompet YSL	12	90	56	34
13	dompet LV	13	45	38	7
57	dompet bally	14	67	67	0
58	dompet GUCCI	15	78	78	0
59	pansus BULAT S	16	90	78	12
60	pansus HOLIC PERMATA besar	17	45	45	0
61	pansus HOLIC silang permata	18	69	67	2
62	pansus pita besar NI	19	78	59	19
63	pansus holic Ceper NI	20	89	49	40

Setelah Data didapatkan, selanjutnya mendeskripsikan data yang bertujuan untuk memahami informasi lebih lanjut mengenai atribut dari data yang akan kita olah.

3.2 Data Cleaning

Merupakan operasi dasar yang dilakukan seperti penghapusan *noise*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data Tahapan Proses KDD [5]. Untuk penelitian ini data selection dilakukan penghapusan pada duplikasi data sebagai berikut:

Tabel 2. Duplikasi data

No	Nama Barang	Kode Barang
1	Sandal ando hawai	01
2	sandal grania nininn NI	02
3	sandal youth balance kning qpr	03
4	sandal karet fimio mntira tbi smg	04
5	sandal youth balance manik cwek	05
6	ransel sekolah import chibao	06
7	ransel santai import	07
8	tas santai lokal	08
9	tas import hermes	09
10	dompet meet you forever	10
11	dompet kartu	11
12	dompet YSL	12
13	dompet LV	13
20	dompet bally	14
21	dompet GUCCI	15
22	pansus BULAT S	16
23	pansus HOLIC PERMATA besar	17
24	pansus HOLIC silang permata	18
25	pansus pita besar NI	19
26	pansus holic Ceper NI	20

Penghapusan data duplikat atau tidak relevan agar tidak berpengaruh pada proses data mining sehingga menjadi data yang sempurna.

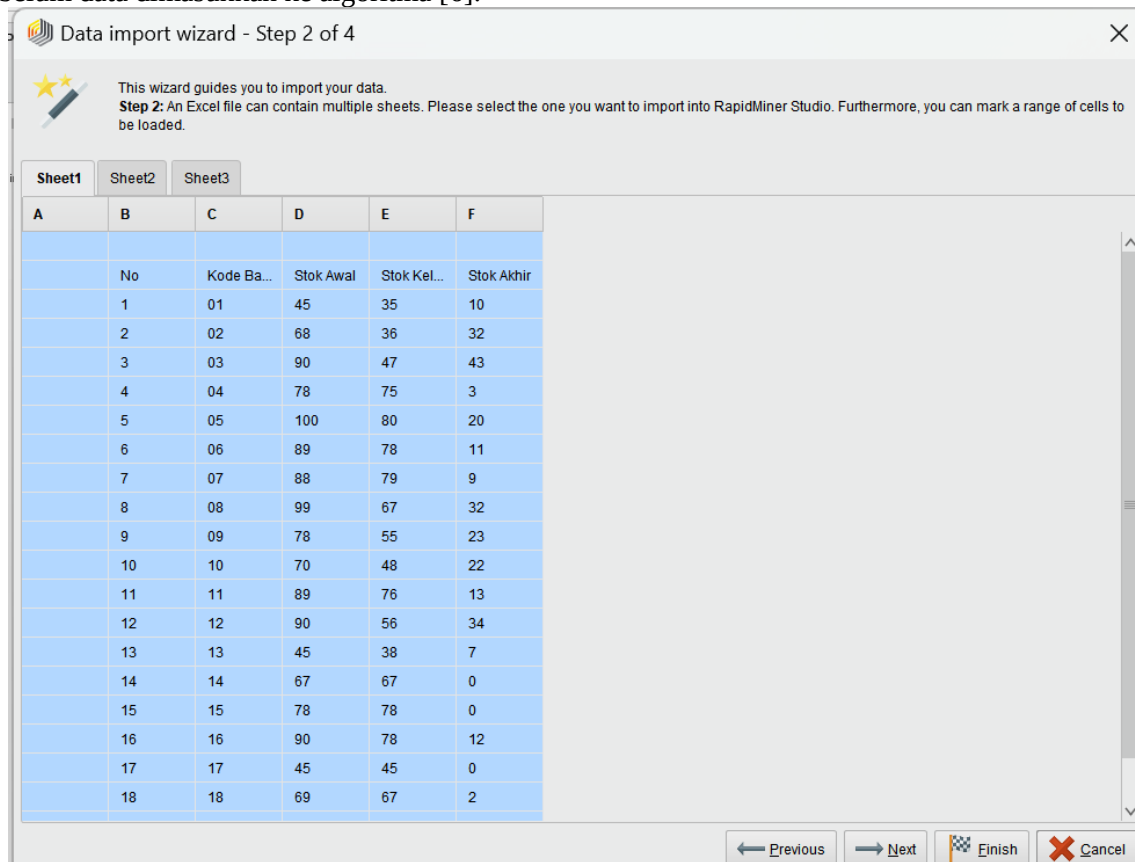
Tabel 3. Hapus Duplikasi Data

No	Nama Barang	Kode Barang
1	Sandal ando hawai	01
2	sandal grania nininn NI	02
3	sandal youth balance kning cpr	03
4	sandal karet fimio mutira tbl smg	04
5	sandal youth balance manik cwek	05
6	ransel sekolah import chibao	06
7	ransel santai import	07
8	tas santai lokal	08
9	tas import hermas	09
10	dompet meet you forever	10
11	dompet kartu	11
12	dompet YSL	12
13	dompet LV	13
14	dompet bally	14
15	dompet GUCCI	15
16	pansus BULAT S	16
17	pansus HOLIC PERMATA besar	17
18	pansus HOLIC silang permata	18
19	pansus pita besar NI	19
20	pansus holic C ceper NI	20

Tabel diatas merupakan hasil pembersihan data duplikat, sehingga mendapatkan data yang sempurna untuk diolah dalam data mining.

3.3 Data Transformasi

Data transformasi mengacu pada proses mengubah atau mempersiapkan data mentah ke dalam format yang sesuai dan optimal untuk proses klasterisasi. Ini merupakan bagian dari tahapan preprocessing sebelum data dimasukkan ke algoritma [6].

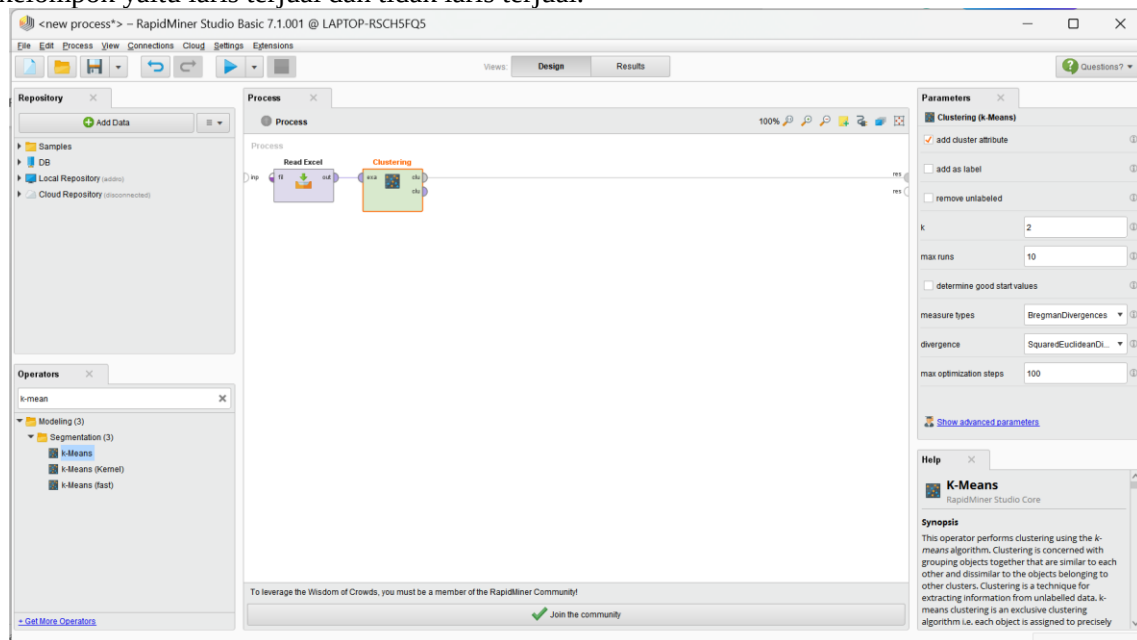


Gambar 3. Data Transformasi

Data transformasi adalah proses mengubah data mentah dari berbagai sumber menjadi informasi yang terstruktur dan siap digunakan untuk analisis bisnis. Di Toko Nusa Indah, data transformasi dilakukan terhadap berbagai data penjualan, seperti transaksi harian, stok barang.

3.4 Data Model

Data mining merupakan proses mencari pola atau informasi menarik dalam data yang terpilih dengan menggunakan metode algoritma *K-Means* [7]. Berdasarkan data gambar 3 tentang model algoritma k-means menjelaskan bahwa operator yang digunakan yaitu retrieve dipergunakan untuk memanggil data set pada penelitian ini. Kemudian clustering k-means untuk memodelkan data set yang telah ada, serta clusterdistance performance digunakan untuk menguji hasil pengelompokan, penelitian ini menetapkan 2 kelompok yaitu laris terjual dan tidak laris terjual.



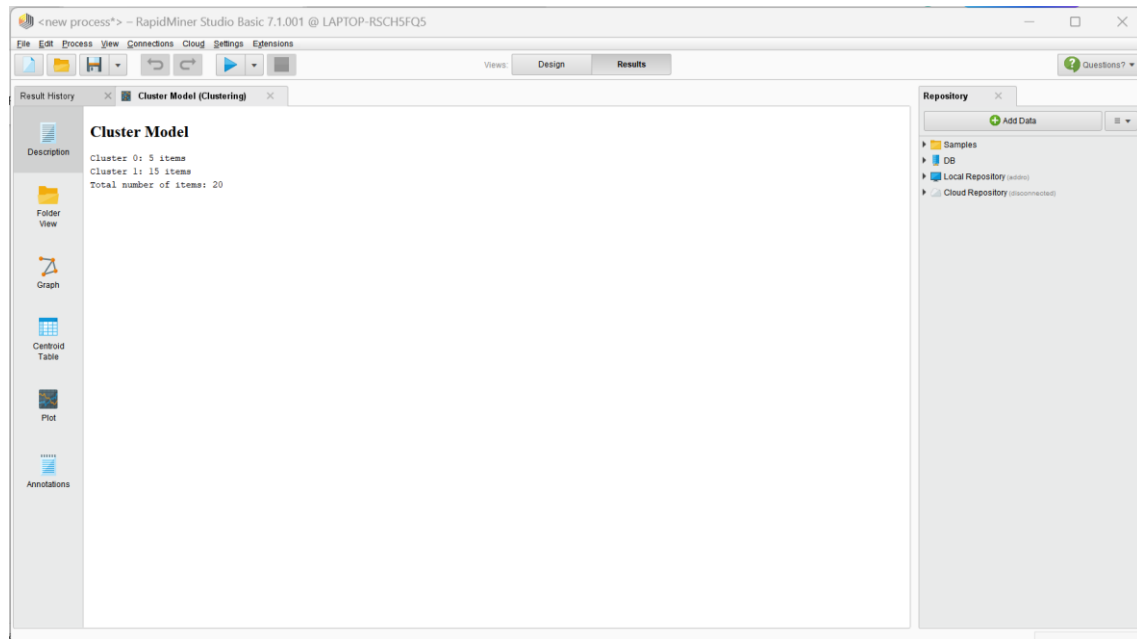
Gambar 4. Model Algoritma K-Means

Gambar diatas merupakan bentuk proses data mining yang menggunakan aplikasi rapid miner. Sehingga data siap untuk diproses untuk menghasilkan *cluster* yang telah kita tentukan.

3.5 Evaluasi

a. Cluster Model

Berdasarkan hasil implementasi Algoritma K-Means maka didapatkan hasil cluster Model sebagai berikut :



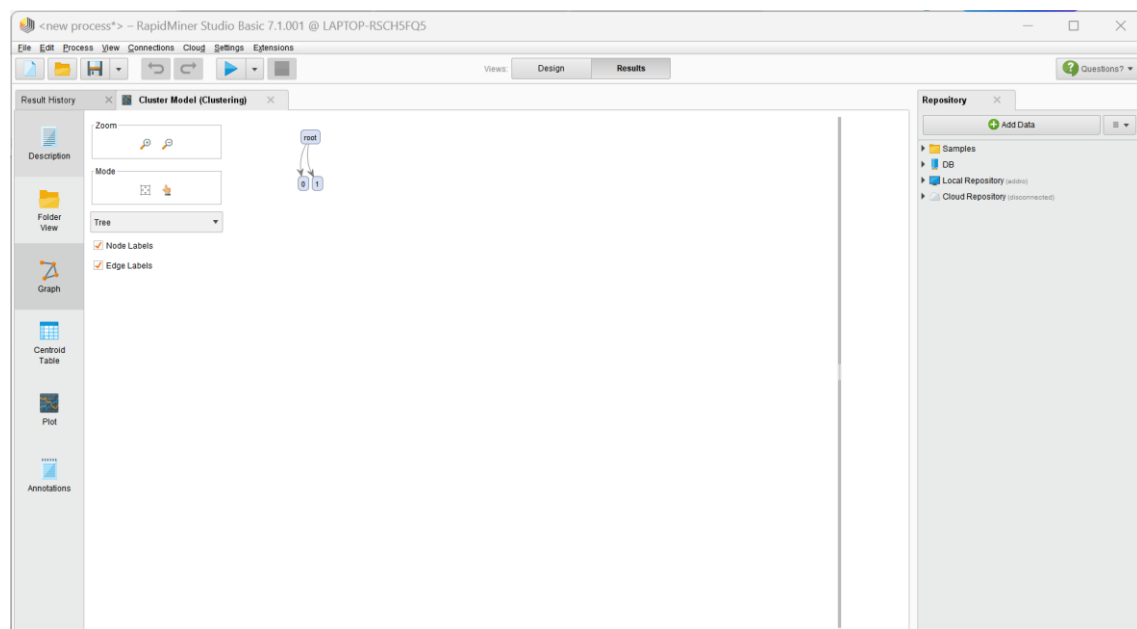
Gambar 5. Hasil Cluster Model

Gambar diatas merupakan hasil cluster model yang telah diproses, cluster 0 memiliki 5 items, dan cluster 1 sebanyak 15 items dimana total item secara keseluruhan yang digunakan saat pengujian sebanyak 20 item data.

b. Graph

Representasi data dan hubungan antar elemen dalam bentuk simpul (nodes) dan sisi (edges).Disajikan pada gambar dibawah ini :

c.

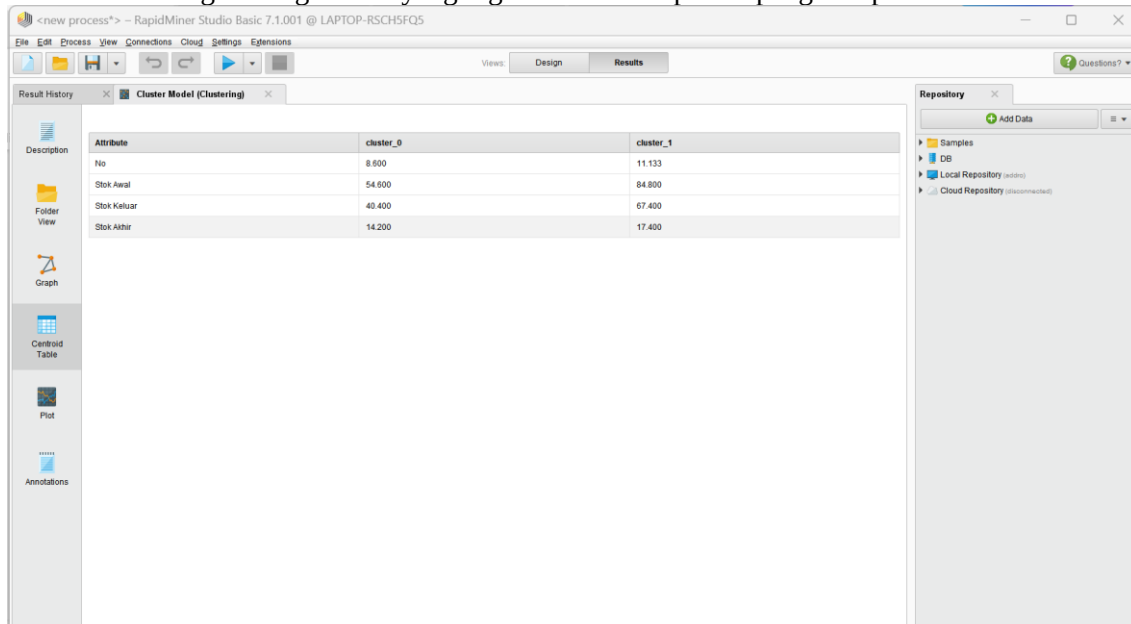


Gambar 6. Graph

Gambar diatas adalah hasil tampilan graph dalam cluster model yang berupa tree.

d. Centroid Table

Tabel centroid di bawah ini menunjukkan hasil dari proses clustering menggunakan metode K-Means, yang menghasilkan 2 klaster (Cluster_0 dan Cluster_1). Setiap cluster memiliki nilai centroid yakni rata-rata dari masing-masing atribut yang digunakan dalam proses pengelompokan data.



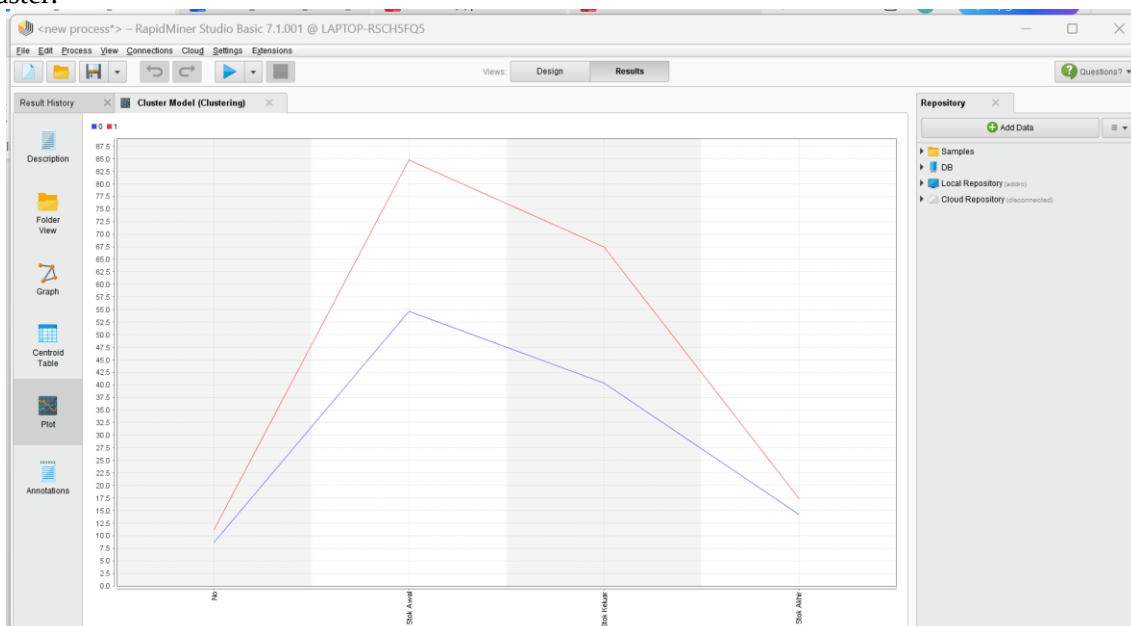
Attribute	cluster_0	cluster_1
No	8.600	11.133
Stok Awal	54.600	84.800
Stok Keluar	40.400	67.400
Stok Akhir	14.200	17.400

Gambar 7. Centroid Table

Gambar diatas merupakan hasil dari proses *clustering* yang menampilkan atribut stok awal, stok keluar, stok akhir. Pada *cluster* 0 stok awal (54.000), stok keluar (40.400), dan stok akhir (14.200). sedangkan *cluster* 1 stok awal (84.400), stok keluar (67.400), dan stok akhir (17.400).

e. Plot

Gambar di bawah ini menampilkan plot visualisasi centroid antar klaster dari hasil proses clustering di RapidMiner Studio. Grafik ini memudahkan kita untuk membandingkan nilai rata-rata tiap atribut antar klaster.



Gambar 8. Plot

Pada gambar diatas plot visualisasi tersebut menampilkan rata-rata setiap *cluster*.

4. KESIMPULAN

Penelitian ini bertujuan untuk menerapkan metode data mining, khususnya algoritma K-Means Clustering, dalam menganalisis data penjualan pada Toko Nusa Indah guna memperoleh wawasan strategis terkait pengelompokan produk berdasarkan performa penjualannya.

Daftar Pustaka

- [1] Pujiono, S., Astuti, R., & Basysyar, F. M. (2024). Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 615-620.
- [2] Afiasari, N., Suarna, N., & Rahaningsi, N. (2023). Implementasi data mining transaksi penjualan menggunakan algoritma clustering dengan metode K-Means. *Jurnal Saintekom: Sains, Teknologi, Komputer dan Manajemen*, 13(1), 100-110.
- [3] Fitri, E. M., Suryono, R. R., & Wantoro, A. (2023). Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz. *Jurnal Komputasi*, 11(2), 157-168.
- [4] Pujiono, S., Astuti, R., & Basysyar, F. M. (2024). Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 615-620.
- [5] Bui, M. A., & Bahtiar, A. (2024). Implementasi metode algoritma K-Means Clustering untuk mengelompokkan transaksi penjualan barang di Toko Arino. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1451-1456.
- [6] Maoulana, R., Irawan, B., & Bahtiar, A. (2024). DATA MINING DALAM KONTEKS TRANSAKSI PENJUALAN HIJAB DENGAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 515- 521.
- [7] Lega, A., Adytia, P., & Lailiyah, S. (2024). Penerapan Algoritma K-means Clustering untuk Klasterisasi Penjualan Smartphone pada Carin Cell. *STMIK Widya Cipta Dharma*.
- [8] Azzahra, L., & Yasir, A. (2024). A Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 3(1), 1-10.
- [9] Widad, R., & Fatah, Z. (2025). Penerapan Algoritma K-Means Clustering Pada Penjualan Produk Skincare. *Jurnal Mahasiswa Teknik Informatika*, 4(1), 52-56.
- [10] Hidayat, K., Adytama, M. R., Darmawan, H. A., Arnando, Y., & Mukarim, A. (2025). Analisis Data Penjualan Menggunakan Algoritma K-Means Clustering Pada Toko Superindo. *JoDMApps (Journal of Data Science Methods and Applications)*, 1(1), 1-6.
- [11] Angraeni, D. F., Rahaningsih, N., Dana, R. D., & Rohmat, C. L. (2025). PEMANFAATAN ALGORITMA K-MEANS DALAM ANALISIS DATA PENJUALAN TOKO BUYUNG UPIK JS DI LAZADA. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(2).
- [12] Yahya, A., & Kurniawan, R. (2025). Implementasi Algoritma K-Means untuk Pengelompokan Data Penjualan Berdasarkan Pola Penjualan: Implementation of K-Means Algorithm for Clustering Sales Data Based on Sales Patterns. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(1), 350-358.
- [13] Agni, M. A. M. (2025). Optimalisasi Pengelompokan Jenis Produk dalam Penentuan Strategi Penjualan Menggunakan Algoritma K-Means di Toko Cirebon Vape Store. *MULTINETICS*, 11(1), 1-10.
- [14] Zou, Z., Ohta, T., & Oki, S. (2024). ChIP-Atlas 3.0: a data-mining suite to explore chromosome architecture together with large-scale regulome data. *Nucleic Acids Research*, 52(W1), W45-W53. Penulis1 A, Penulis2 B. Judul Buku. Edisi. Kota: Penerbit. tahun: halaman.
- [15] Tsiu, S. V., Ngoben, M., Mathabela, L., & Thango, B. (2025). Applications and competitive advantages of data mining and business intelligence in SMEs performance: A systematic review. *Businesses*, 5(2), 22.
- [16] Zhou, X., Du, H., Xue, S., & Ma, Z. (2024). Recent advances in data mining and machine learning for enhanced building energy management. *Energy*, 132636.

- [17] Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46-56.
- [18] Manalu, D. A., & Gunadi, G. (2022). Implementasi Metode Data Mining K-Means Clustering Terhadap Data Pembayaran Transaksi Menggunakan Bahasa Pemrograman Python Pada Cv Digital Dimensi. *Infotech: Journal of Technology Information*, 8(1), 43-54.
- [19] Nugraha, A., Nurdiawan, O., & Dwilestari, G. (2022). Penerapan data mining metode K-Means clustering untuk analisa penjualan pada toko Yana Sport. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 849-855.
- [20] Afiasari, N., Suarna, N., & Rahaningsi, N. (2023). Implementasi data mining transaksi penjualan menggunakan algoritma clustering dengan metode K-Means. *Jurnal Saintekom: Sains, Teknologi, Komputer dan Manajemen*, 13(1), 100-110.
- [21] Kurniawan, H., & Defit, S. (2020). Data mining menggunakan metode k-means clustering untuk menentukan besaran uang kuliah tunggal. *Journal of Applied Computer Science and Technology*, 1(2), 80-89.
- [22] Fauzi, I. (2025). Penerapan Data Mining Untuk Menentukan Lokasi Promosi Sekolah Dengan Metode K-Means Clustering:(Studi Kasus: SMP Islam Al Syukro Universal). *Jurnal Informatika Utama*, 3(1), 112-126.
- [23] Irawan, D., Wijaya, G., & Warisaji, T. T. (2025). Penerapan Algoritma K-Means Clustering untuk Segmentasi Nasabah Bank. *BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer*, 6(1), 47-53.
- [24] Sari, R. A. K. N., & Kurniawati, D. (2025). PENERAPAN DATA MINING DALAM ANALISIS POLA PENJUALAN PRODUK CETAKAN DENGAN K-MEANS CLUSTERING. *Journal of Innovation Research and Knowledge*, 5(1), 435-446.
- [25] Nasution, N. W., Ilmi, N., & Candra, J. E. (2025). IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM PENGELOMPOKAN HASIL TANGKAP IKAN DI KEPULAUAN RIAU. *Journal of Innovation Research and Knowledge*, 5(1), 693-704.



Startup: Journal of Digital Business

is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)