

Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam

Gita Aprilianur¹⁾, Elvin Leander Hadisaputro²⁾

^{1,2*)} Program Studi Sistem Informasi, STMIK Borneo Internasional Balikpapan, Balikpapan.
Jalan AW Syahrane, Batu ampar, Kota Balikpapan, Kalimantan Timur 76136

e-mail: ¹⁾ gita_aprilianur.18@stmik-borneo.ac.id, ²⁾ elvin.leander@stmik-borneo.ac.id

Abstrak

Toko Myam Hijab merupakan toko yang khusus menjual hijab, namun dari berbagai hijab yang dijual tentunya tidak semuanya laku, ada juga yang tidak laku. Data penjualan, pembelian barang dan pengeluaran tak terduga di Myam Hijab tidak terstruktur dengan baik, sehingga data tersebut hanya berfungsi sebagai arsip toko dan tidak dapat digunakan untuk mengembangkan strategi pemasaran. Oleh karena itu, perlu diterapkan data mining dengan menggunakan metode K-Means. di toko Myam Hijab. Metode K-Means dapat diterapkan pada Myam Hijab Store untuk mengetahui penjualan hijab mana yang laris, laris dan tidak laris. Penerapan metode K-Means pada toko Myam ini yaitu dengan mengelompokkan data stok hijab. Kemudian pilih 3 grup acak sebagai centroid awal. Setelah data tiap kelompok tidak berubah, terlihat bahwa hasil akhirnya ada 24 produk yang laku keras, 59 produk laris, dan 17 produk kurang laku. Kemudian, penerapan metode K-Means pada Rapidminer dilakukan dengan memasukkan data stok produk yaitu stok awal, stok terjual dan stok akhir, yang akan diubah menjadi database di Ms.Excel, data tersebut akan terhubung ke Rapidminer Tools dan diproses. dan membentuk K-means. Setelah itu, Rapidminer akan memproduksi produk mana yang permintaannya tinggi, permintaannya tinggi, dan permintaannya rendah.

Kata kunci— Data Mining, K-Means, Myam Hijab dan Rapidminer.

Abstract

Myam Hijab shop is a store that specializes in selling hijab, but of the various hijabs that are sold, of course, not all of them sell, some do not. Sales data, purchases of goods and unexpected expenses in Myam Hijab are not well structured, so the data only serves as a store archive and cannot be used to develop marketing strategies. Therefore, it is necessary to apply data mining using the K-Means method. at the Myam Hijab shop. The K-Means method can be applied to the Myam Hijab Store to find out which hijab sales are selling, selling and not selling. The application of the K-Means method at this Myam shop is by grouping hijab stock data. Then select 3 random groups as the initial centroid. After the data for each group did not change, it was seen that in the end there were 24 products that were selling well, 59 products that were selling well, and 17 products that were not selling well. Then, the application of the K-Means method on Rapidminer is done by entering product stock data, namely initial stock, sold stock and ending stock, which will be converted into a database in Ms.Excel, the data will be connected to Rapidminer Tools and processed. and form K-means. After that, Rapidminer will produce which products are in high demand, high demand, and low demand.

Keywords— Data Mining, K-Means, Myam Hijab and Rapidminer.

1. PENDAHULUAN

Dalam dunia bisnis yang kompetitif saat ini, masyarakat dituntut untuk selalu mengembangkan usahanya agar dapat bertahan dalam persaingan terutama dalam persaingan penjualan, yang menuntut para pengusaha untuk menemukan pola-pola yang dapat meningkatkan penjualan dan pemasaran di dalam perusahaan, salah satunya adalah pemanfaatan penjualan [1]. Tidak digunakan oleh salah satu toko hijab yaitu Myam Hijab Shop. Myam Hijab Shop adalah toko yang didedikasikan untuk menjual hijab, namun tentunya tidak semua berbagai jenis hijab itu sangat laris, laris, kurang laris. Data penjualan, pembelian barang dan pengeluaran tak terduga di toko Myam Hijab tidak tertata dengan baik, sehingga data tersebut hanya berfungsi sebagai arsip toko dan tidak dapat digunakan untuk pengembangan strategi pemasaran. Data yang tersedia seharusnya dapat digunakan sebagai sistem pengambilan keputusan untuk solusi bisnis dan dukungan infrastruktur di bidang teknologi yang menjadi penyebab munculnya sebuah teknologi yaitu data mining [1]

Data mining adalah proses penggalian data dari informasi yang sangat penting. Data mining juga merupakan proses eksplorasi pola dari data. Pola diperoleh dari berbagai jenis database, seperti database relasional, data warehouse, data transaksional, dan data berorientasi objek. Penggunaan data mining dapat membantu pengusaha mengambil keputusan dengan cepat dan akurat.

K-Means adalah metode clustering data non-hierarchical yang berusaha untuk mempartisi data yang ada menjadi satu atau lebih cluster atau Group sehingga data dengan karakteristik yang sama dikelompokkan ke dalam cluster yang sama [2]. K-Means merupakan metode clustering berbasis jarak yang membagi data menjadi beberapa cluster dan algoritma ini hanya bekerja dengan atribut numerik [3]. Menentukan pola penjualan dengan melihat tren pembelian konsumen, jika dianalisa dan diolah dengan baik, dapat membantu menentukan produk mana yang paling laris dan produk mana yang paling sedikit penjualannya, sehingga dapat membawa inventaris Anda dan dapat Anda gunakan sebagai bahan masukan bagi perusahaan yang mengembangkan strategi pemasaran [4].

Output data mining dapat digunakan untuk membantu pengambilan keputusan di masa depan. "K-Means adalah metode clustering non-hierarchical yang mencoba untuk mempartisi data yang ada menjadi satu atau lebih cluster atau mungkin bertujuan untuk mempartisi data menjadi beberapa cluster [5]." Algoritma k-means digunakan untuk mengetahui minat pelanggan terhadap produk yang dijual [2]. Tujuan dari penelitian ini adalah untuk membantu Myam Hijab Shop menentukan produk jual yang termasuk dalam kategori sangat laris, laris dan kurang laris.

2. METODE PENELITIAN

Dalam penelitian ini dilakukan beberapa tahapan secara sistematis, dengan tujuan untuk memudahkan peneliti dalam mengklasifikasikan prosedur dan urutan dalam suatu penelitian dan lebih terarah, penelitian ini merupakan penelitian investigasi yang menggunakan model penelitian kuantitatif [6]. Model kuantitatif ini membantu dan membimbing dalam suatu penyelidikan berupa pemecahan suatu masalah dalam bentuk penelitian agar pencapaian tujuan penyelidikan lebih on track. Dalam penelitian ini, data dikumpulkan dengan menggunakan berbagai teknik pengumpulan data sebagai berikut:

1. Observasi
Penulis melakukan observasi langsung di bagian penjualan toko hijab myam yang berlokasi di Jalan Raden Sukma RT. 21 Penajam Paser Utara.
2. Wawancara
Penulis memperoleh data dan meneliti kebenaran informasi pada data ini dengan melakukan tanya jawab langsung dengan pemilik toko, Ibu Anah.

3. Studi Pustaka

Melakukan penelitian kepustakaan untuk mengumpulkan data dan informasi yang diperoleh dari perpustakaan, jurnal, artikel, dan internet yang berhubungan dengan penelitian ini.

Kerangka penelitian ini dijelaskan di bawah ini:



Gambar 1. Kerangka kerja penelitian

Tahapan penelitian ini dapat digambarkan sebagai berikut:

1. Identifikasi masalah
Yaitu melakukan permasalahan yang terdapat di toko yaitu untuk mengetahui produk mana yang banyak diminati, diminati dan tidak diminati sebagai strategi utama penyediaan stok hijab di toko.
2. Pengumpulan data
Mengumpulkan data yaitu meminta data stok hijab berdasarkan data stok hijab periode Oktober 2021 sampai dengan Desember 2021 di toko Myam Hijab.
3. Analisis data
Yaitu menganalisis masalah yaitu menganalisis data stok hijab di toko Myam Hijab.
4. Pemrosesan data
Yaitu dengan mengolah data-data yang diberikan oleh toko Myam Hijab, kemudian data tersebut diolah dengan menggunakan metode K-means clustering yang dihitung secara manual dengan untuk mengklasifikasikan produk yang diminati, diminati dan tidak diminati.
5. Uji data
Yaitu dengan menguji data yang sudah diolah dengan Rapid Miner.
6. Kesimpulan
Yaitu penarikan kesimpulan dari hasil perhitungan dan pengolahan data yang dihitung secara manual dengan data Rapid Miner.

Data yang digunakan dalam penelitian ini adalah data stok hijab, tas, cardy pada bulan Oktober 2021 – Desember 2021 di toko myam hijab [1]. Berikut data-data yang akan diolah, berupa stok awal, stok terjual dan stok akhir:

Tabel 1. Data Stok Toko Myam Hijab Pada Bulan Oktober 2021 – Desember 2021

No	Produk	SA	ST	SAK
1	Pashmina Chayra Abu	55	20	35
2	Pashmina Chayra lavender	55	15	40
3	Pashmina Chayra Milo	35	10	25
4	Pashmina Chayra Navy	40	15	25
5	Pashmina Chayra Light Grey	35	20	15
6	Pashmina Chayra Latte	45	35	10
7	Pashmina Chayra Rassberry	50	40	10
8	Pashmina Chayra Dove	50	35	15
9	Pashmina Chayra Mocca	35	30	5
10	Pashmina Chayra Tailo	35	20	15
11	Pashmina Chayra Broken white	45	25	20
12	Pashmina Chayra Coklat Pasta	45	20	25
13	Pashmina Chayra Coklat Grand	45	15	30
14	Pashmina Chayra Army	35	25	10
15	Pashmina Chayra Maroon	30	20	10
16	Pashmina Chayra Lime	30	10	20
17	Pashmina Chayra Dusty Pink	30	5	25
19	Pashmina Chayra dark grey	45	20	25
20	Pashmina Chayra Hitam	35	25	10
21	Scrunchie Maroon	20	5	15
22	Scrunchie Dusty Pink	20	10	10
23	Scrunchie Lime	20	15	5
24	Scrunchie Army	20	4	16
25	Scrunchie Blue	20	5	15
26	Scrunchie Coklat	20	10	10
27	Scrunchie Milo	20	15	5
28	Scrunchie Grey	20	2	18
29	Scrunchie Hitam	20	18	2
30	Scrunchie Broken White	20	13	7
31	Scrunchie Lilac	20	17	3
32	Tas Merble Hitam	10	4	6
33	Tas Merble Cream	10	3	7
34	Tas Merble Pink	10	5	5
35	Tas Merble Putih	10	7	3
36	Dompot Panjang Hitam	15	4	11
37	Dompot Panjang Blue	10	3	7
38	Dompot Panjang Abu	10	7	3
39	Dompot Bunga	15	9	6
40	Dompot Kelinci	10	3	7
41	Piyama Tyedie Army	10	4	6
42	Piyama Tyedie Navy	15	10	5
43	Piyama Tyedie Brown	10	7	3
44	Piyama Tyedie Abu	15	12	3
45	Piyama Tyedie Hitam	15	13	2

46	Segiempat Nur Dusty Pink	35	15	20
47	Segiempat Nur Mocca	35	20	15
48	Segiempat Nur Army	40	28	12
49	Segiempat Nur Navy	40	37	3
50	Segiempat Nur Abu	40	25	15
51	Segiempat Nur Milo	40	24	16
52	Segiempat Nur Lime	45	25	20
53	Segiempat Nur Maroon	45	35	10
54	Segiempat Nur Cream	50	45	5
55	Segiempat Nur Hitam	50	48	2
56	Cardy Rajut Abu	5	1	4
57	Cardy Rajut Denim	5	2	3
58	Square Naura syari Hitam	20	16	4
59	Square Naura syari Dark Grey	20	12	8
60	Square Naura syari Broken White	25	15	10
61	Square Naura syari Abu	25	10	15
62	Square Naura syari Light Grey	20	13	7
63	Square Naura syari Army	10	4	6
64	Square Naura syari Navy	10	7	3
65	Square Naura syari Coklat Pasta	25	15	10
66	Square Naura syari Baby Pink	25	10	15
67	Square Naura syari Dusty Pink	20	10	10
68	Square Naura syari Lavender	20	15	5
69	Square Naura syari Rassbery	20	18	2
70	Square Naura syari Mocca	25	22	3
71	Square Naura syari Tailo	25	10	15
72	Square Naura syari Maroon	10	6	4
73	Khimar Instan Haura Milo	10	3	7
74	Khimar Instan Haura Purple	25	20	5
75	Khimar Instan Haura Hitam	35	10	25
76	Khimar Instan Haura Mocca	45	30	15
77	Khimar Instan Haura Coklat susu	10	4	6
78	Khimar Instan Haura Maroon	10	2	8
79	Khimar Instan Haura Lavender	25	12	13
80	Khimar Instan Haura Dark Grey	40	28	12
81	Khimar Instan Hauri Milo	45	25	20
82	Khimar Instan Hauri Purple Ash	35	10	25
83	Khimar Instan Hauri White	45	30	15
84	Khimar Instan Hauri Dark Grey	20	10	10
85	Khimar Instan Hauri Hitam	45	35	10
86	Khimar Instan Hauri Mocca	10	7	3
87	Khimar Instan Hauri Coklat Susu	10	5	5
88	Khimar Instan Hauri Milk Choco	15	10	5
89	Khimar Instan Hulya Hitam	20	10	10
90	Khimar Instan Hulya Lavender	25	15	10
91	Khimar Instan Hulya Mocca	45	10	35
92	Khimar Instan Hulya Dusty Lilac	50	34	16
93	Khimar Instan Hulya White	25	15	10
94	Khimar Instan Hulya Maroon	45	25	20
95	Khimar Instan Hulya Baby Pink	30	10	20
96	Khimar Instan Hulya Milo	10	5	5

97	Khimar Instan Hulya Dark Grey	10	8	2
98	Lulur Bedak Lotong 200kg	15	9	6
99	Lulur Bedak Lotong 1 kg	10	5	5
100	Gift Box	20	15	5

Langkah-langkah dalam mengcluster menggunakan metode K-Means adalah sebagai berikut :

1. Tentukan nilai k nya sebagai jumlah kluster yang akan dibentuk.
2. Tentukan Titik pusat awal dari setiap kluster.
3. Hitunglah jarak setiap data input masing – masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) sampai ditemukan jarak yang terdekat dari setiap data dengan centroid.

Berikut adalah persamaan Euclidian Distance :

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

Keterangan;

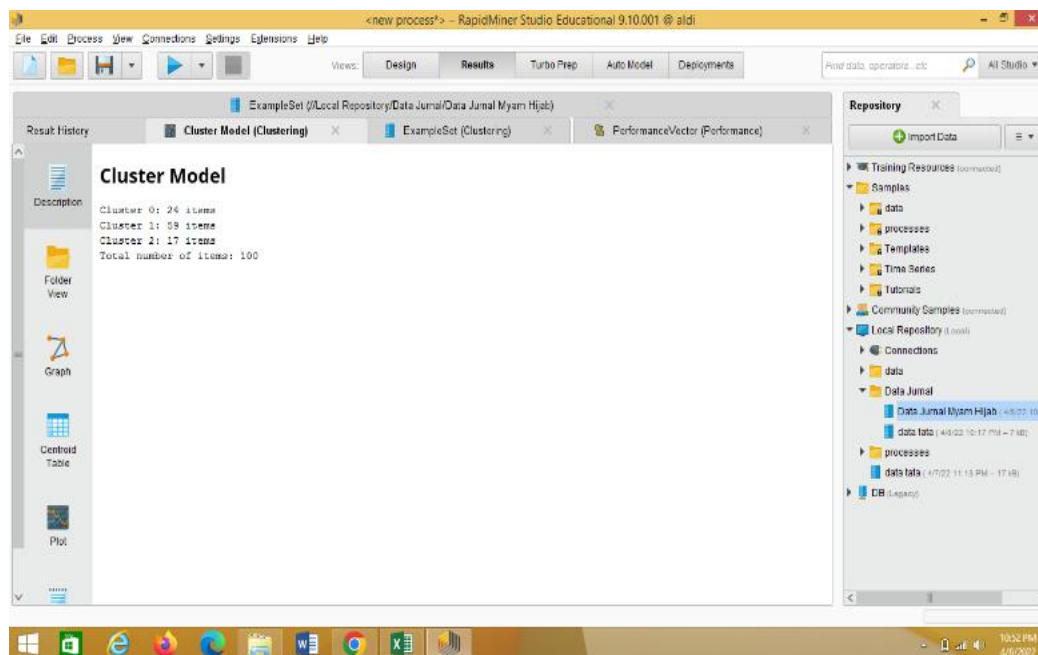
D = Jarak x = Data

y = Centroid

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.
6. 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.

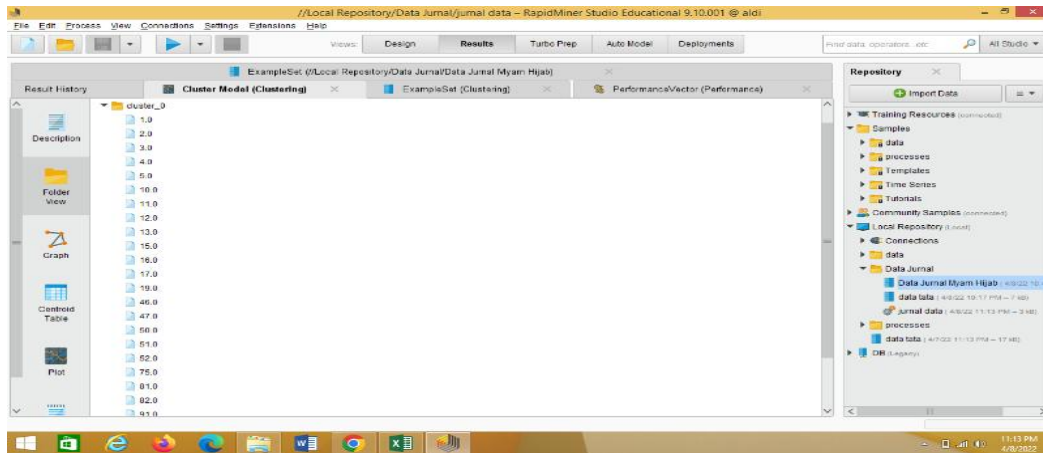
3. HASIL DAN PEMBAHASAN

1. Tampilan Hasil Clustering
- 2.



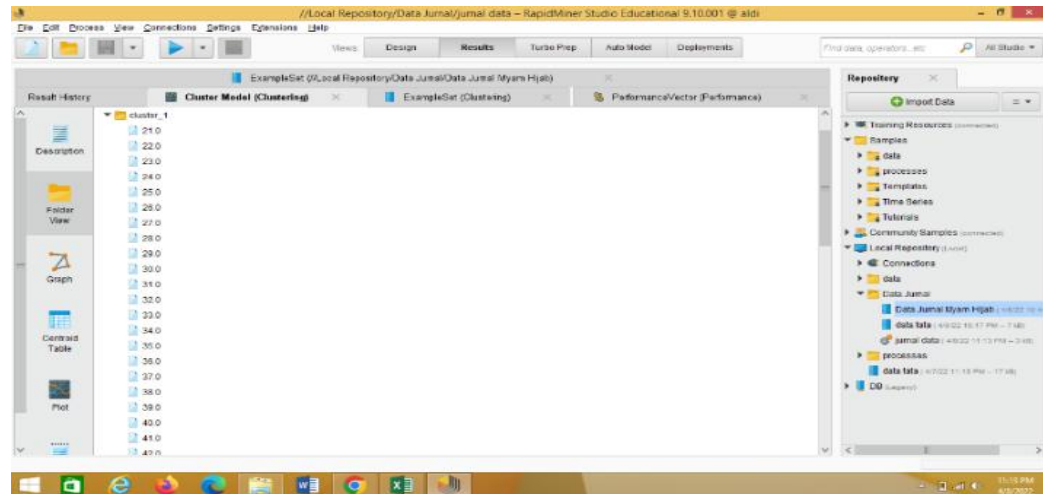
Gambar 2. Tampilan Hasil Cluster Model

Gambar 2 adalah tampilan hasil cluster model dari data yang diolah.



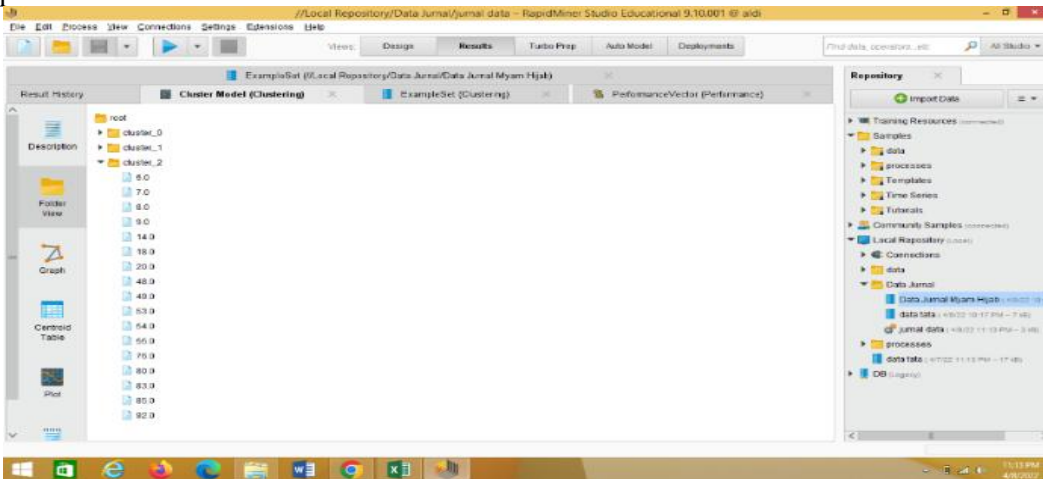
Gambar 3. Tampilan Hasil Cluster 1

Gambar 3 adalah tampilan hasil cluster 1 yang sudah diolah dari Rapidminer yang terdapat 24 produk data.



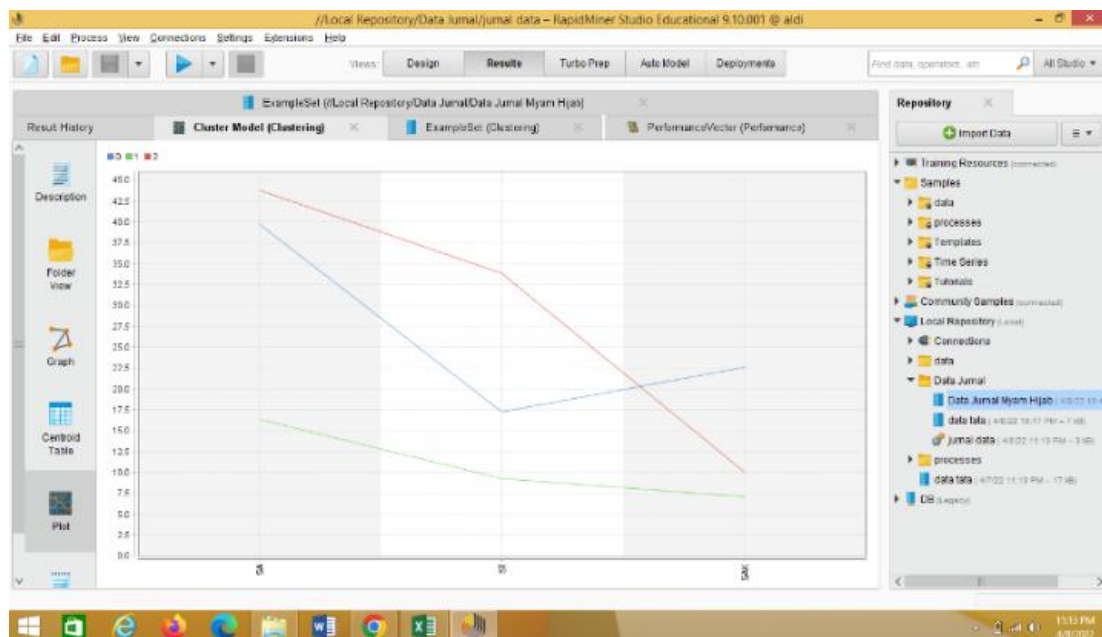
Gambar 4. Tampilan Hasil Cluster 2

Gambar 4 adalah tampilan hasil dari cluster 2 yang sudah diolah dari Rapidminer yang terdapat 59 produk data.



Gambar 5. Tampilan Hasil Cluster 3

Gambar 5 adalah tampilan hasil dari cluster 3 yang sudah diolah dari Rapidminer yang terdapat 17 produk data.



Gambar 6. Hasil Pengolahan Data Pada Rapidminer

Gambar 6 merupakan hasil pengolahan data pada Rapidminer yang ditandai dengan warna biru merupakan hasil dari cluster 1, warna hijau adalah hasil dari cluster 2 dan warna merah merupakan hasil dari cluster 3.

Tabel 2. Centroid Awal

Keterangan	SA	ST	SAK
C1	5	1	4
C2	35	10	25
C3	45	15	30

Tabel 2 adalah Menentukan Pusat Awal Cluster secara acak diambil dari data pada Tabel 1. Data yang dipilih untuk cluster awal adalah produk ke-3, 13, 56. Jarak centroid data ke-1 pada cluster 1 (C1) yaitu:

Cluster 1

$$\sqrt{(SA_1 - C1_1)^2 + (ST_2 - C2_2)^2 + (SAK_3 - C3_3)^2} = \dots$$

$$\sqrt{(55_1 - 5_1)^2 + (20_2 - 1_2)^2 + (35_3 - 4_3)^2} = 37.0405$$

Cluster 2

$$\sqrt{(55_1 - 35_1)^2 + (20_2 - 10_2)^2 + (35_3 - 25_3)^2} = 14.8323$$

Cluster 3

$$\sqrt{(55_1 - 45_1)^2 + (20_2 - 15_2)^2 + (35_3 - 30_3)^2} = 7.7459$$

4. KESIMPULAN

Sehingga dari hasil penelitian dapat disimpulkan bahwa Metode K-means dapat diterapkan pada toko Myam Hijab untuk menentukan penjualan baju mana yang sangat laris, laris dan kurang laris. Penerapan metode K-Means pada Myam Hijab yaitu dengan cara mengelompokan data stok hijab. Kemudian memilih 3 cluster secara acak sebagai centroid awal. Setelah data pada setiap *cluster* tidak berubah-ubah, maka dapat diketahui hasil akhirnya yaitu yang sangat laris ada 24 produk, yang laris ada 59 produk dan 17 produk untuk yang kurang laris. Menerapkan metode *K-means* pada Rapidminer dilakukan dengan memasukkan data stok produk yaitu stok awal, stok terjual dan stok akhir yang akan menjadi *Database* pada Ms.Excel, data tersebut kemudian dikoneksikan ke dalam *Tools* Rapidminer, dan akan diolah dan dibentuk *K-means*. Setelah itu Rapidminer akan menghasilkan produk mana yang sangat laris, laris, dan kurang laris.

5. SARAN

Pada penelitian ini masih memiliki banyak kekurangan sehingga diperlukan rencana dalam perbaikan seperti Penelitian ini hanya menggunakan data penjualan selama 3 bulan yaitu pada bulan oktober 2021 sampai dengan desember 2021.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ketua STMIK Borneo Internasional Balikpapan karena telah memberikan kesempatan untuk melakukan penelitian Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam.

DAFTAR PUSTAKA

- [1] I. Wahyudi, S. Bahri, and P. Handayani, "Aplikasi Pembelajaran Pengenalan Budaya Indonesia," vol. V, no. 1, pp. 135–138, 2019.
- [2] Y. Yulianti, D. Y. Utami, N. Hikmah, and F. N. Hasan, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 241–246, 2019.
- [3] B. M. Metisen and H. L. Sari, "Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila," *J. Media Infotama*, vol. 11, no. 2, pp. 110–118, 2015.
- [4] N. Normah, B. Rifai, and P. Sari, "Algoritma Apriori Sebagai Solusi Kontrol Persediaan Suku Cadang Mobil PT. Buanasakti Aneka Motor Jakarta," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 2, pp. 161–168, 2020.
- [5] A. Bahar, B. Pramono, and L. H. S. Sagala, "Penentuan strategi penjualan alat-alat tattoo di studio sonyxtattoo menggunakan metode," *semanTIK*, vol. 2, no. 2, pp. 75–86, 2016.
- [6] B. G. Sudarsono and S. P. Lestari, "Clustering Penerima Beasiswa Yayasan Untuk Mahasiswa Menggunakan Metode K-Means," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 258, 2021.
- [7] A. M. Sari, "Aplikasi Situs Web Penjualan Hijab Online," *Paradigma*, vol. 19, no. 2, pp. 113–117, 2017.
- [8] A. Nur istiani, "Kontruksi Makna Hijab Fasion Bagi Muslim Blogger Muslim," *J. Kaji. Komunikas*, vol. 3, no. 1, pp. 48–55, 2015.
- [9] I. P. E.- Issn, "Analisis Clustering Menggunakan Algoritma K-Means Terhadap

- Penjualan Produk Padapt Batamas Niaga Jaya,” *Comput. Based Inf. Syst. J.*, vol. 02, pp. 20–35, 2018.
- [10] F. Rahman, D. Muhammad, and I. Firdaus, “Penerapan Data Mining Metode Naïve Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah Menengah Pertama (Smp),” *Al Ulum Sains dan Teknol.*, vol. 1, no. 2, pp. 76–78, 2016.
- [11] S. Haryati, A. Sudarsono, and E. Suryana, “Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu),” *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2015.