

Hybrid Data Mining berdasarkan Klusterisasi Produk untuk Klasifikasi Penjualan

Dewi Eka Putri^{1✉}, Eka Praja Wiyata Mandala²

^{1,2} Universitas Putra Indonesia YPTK Padang

dewieka@upiptyk.ac.id

Abstract

Sales at minimarkets have had ups and downs since the Covid-19 pandemic. In this study, the problem was found, the number of products in minimarkets that were not sold or there was a demand for products, but they were not available in minimarkets. Data mining can be a solution in solving this problem. This study proposes a hybrid data mining method by combining the K-Means and K-Nearest Neighbors algorithms. The combination of this method works in two stages, clustering the products sold and then continued by classifying the sales of these products. The data used in this study is data for toiletries and washing products as many as 20 products. From the results of this study, there are 14 products that have Many Enthusiasts from 20 products. Of the 14 products used as training data, a sales classification was carried out for 1 new product with stock criteria 200, sold 100 and price Rp. 10,000. From the test results, the new product classification is High Sales with accuracy in the classification reaching 85.7143%. The use of a hybrid method between the K-Means and K-Nearest Neighbors algorithms has a significant influence in determining the classification results.

Keywords: Hybrid Data Mining, Clustering, Classification, Sales, Product

Abstrak

Penjualan yang terjadi pada minimarket mengalami pasang surut sejak pandemi Covid-19. Dalam penelitian ini, didapatkan masalah yaitu banyaknya produk pada minimarket tidak laku terjual atau sebaliknya, adanya permintaan produk, namun di minimarket tidak tersedia. Data mining bisa menjadi solusi dalam menyelesaikan masalah ini. Penelitian ini mengusulkan metode *hybrid data mining* dengan menggabungkan algoritma *K-Means* dan *K-Nearest Neighbors*. Kombinasi metode ini bekerja dengan dua tahap yaitu melakukan klusterisasi produk yang terjual yang diteruskan dengan melakukan klasifikasi penjualan produk tersebut. Data yang digunakan pada penelitian ini adalah data produk perlengkapan mandi dan mencuci sebanyak 20 produk. Dari hasil penelitian yang dilakukan, terdapat 14 produk yang banyak peminat dari 20 produk. Dari 14 produk yang digunakan sebagai data pelatihan, dilakukan klasifikasi penjualan untuk 1 produk baru dengan kriteria stok ≥ 200 , terjual ≥ 100 dan harga \geq Rp. 10.000. Dari hasil pengujian diperoleh klasifikasi produk baru adalah penjualan tinggi dengan ketepatan dalam klasifikasi mencapai 85.7143 %. Penggunaan metode *hybrid* antara algoritma *K-Means* dan *K-Nearest Neighbors* memberikan pengaruh yang signifikan dalam menentukan hasil klasifikasi.

Kata kunci: Hybrid Data Mining, Klusterisasi, Klasifikasi, Penjualan, Produk

KomtekInfo is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



1. Pendahuluan

Keberadaan minimarket saat ini sudah mengalahkan warung tradisional karena lengkapnya produk yang dijual pada minimarket tersebut. Minimarket merupakan sejenis pasar dimana pembelinya dapat melayani dirinya sendiri dengan hanya mempunyai luas 300 m² sampai 500 m² yang memiliki produk yang cukup lengkap untuk dijual [1]. Minimarket menjual produk yang hampir sama dengan warung tradisional [2], sehingga membuat masyarakat merasa dimudahkan dalam mencari keperluan yang dibutuhkan.

Namun, bagi minimarket hal ini menjadi masalah tersendiri. Dengan tingginya tingkat penjualan yang terjadi, maka pihak minimarket menghadapi masalah tentang pengelompokan produk yang dijual. Akibatnya, akan berdampak pada penjualan minimarket tersebut. Masalah yang paling sering dialami oleh

minimarket adalah terjadinya penumpukan produk yang tidak laku terjual [3]. Masalah lainnya adalah permintaan produk yang banyak namun stok yang tersedia tidak mencukupi [4].

Untuk mengatasi masalah tersebut, maka pada penelitian ini digunakan pendekatan *data mining*. *Data mining* merupakan teknik yang digunakan untuk menghitung dengan menggabungkan statistik, matematika dan *machine learning* [5]. Perkembangan *data mining* saat ini sudah sangat banyak membantu para pengambil keputusan untuk menentukan keputusan yang tepat terhadap masalah yang dihadapi [6].

Penelitian ini mengusulkan penggabungan dua metode dalam *data mining* yaitu klusterisasi dan klasifikasi. Penggabungan metode *data mining* ini disebut dengan *hybrid data mining*. *Hybrid data mining* merupakan kombinasi dari beberapa metode fitur seleksi dan klasifikasi untuk membantu proses pengambilan

keputusan [7]. Dalam metode *hybrid data mining*, faktor penting yang mempengaruhi adalah proses pembangkitan berbagai klasifikasi untuk membuat model umum. Ada banyak cara berbeda untuk menghasilkan pengklasifikasi yang berbeda dan bervariasi [8].

Penelitian ini dimulai dengan melakukan klusterisasi produk yang dijual pada minimarket. Klusterisasi merupakan salah satu metode yang memiliki fungsi penyelesaian masalah dalam *data mining* [9]. Klusterisasi adalah suatu metode yang sering digunakan untuk melakukan pengelompokan berdasarkan kemiripan [10]. Klusterisasi produk ini dilakukan dengan menggunakan algoritma *K-Means*. Algoritma *K-Means* merupakan salah satu algoritma yang digunakan dalam proses klusterisasi [11]. *K-Means* dapat membantu dalam mengelompokkan data penjualan produk untuk menjaga persediaan produk dan mengantisipasi tinggi permintaan pelanggan [12].

Setelah dilakukan pengelompokan produk, penelitian ini dilanjutkan dengan melakukan klasifikasi untuk mendapatkan tingkat penjualan produk tersebut. Klasifikasi merupakan salah satu metode dalam *data mining* yang digunakan untuk membagi atau mengelompokkan data [13] dan menemukan pola untuk memisahkan kelas data [14]. Proses klasifikasi penjualan produk ini dilakukan dengan menggunakan algoritma *K-Nearest Neighbors* (K-NN). K-NN termasuk ke dalam algoritma pembelajaran yang diawasi dan memiliki tingkat akurasi yang tinggi [15]. K-NN digunakan untuk melakukan klasifikasi produk berdasarkan jarak yang paling dekat dengan kelompok produk [16].

Penelitian sebelumnya tentang klusterisasi data penjualan produk pada minimarket menggunakan hanya algoritma *K-Means*. Penelitian ini dilakukan di minimarket Berkah Abadi dengan menggunakan 278480 data transaksi yang dikelompokkan ke dalam 3 *cluster* penjualan, yaitu terbanyak, sedang dan rendah. Penelitian ini menghasilkan tingkat akurasi dengan *confusion matrix* mencapai 87% [17]. Penelitian lain tentang klusterisasi dengan algoritma *K-Means* dilakukan pada Alfamart Cikarang. Penelitian ini menggunakan data sebanyak 67 produk yang dibagi menjadi 2 *cluster* yaitu *Slow Moving Products* dan *Fast Moving Products*. Hasil penelitian ini menjelaskan bahwa dengan dilakukannya kluster produk, maka akan membantu Alfamart Cikarang dalam melakukan prediksi stok produk kedepannya [18].

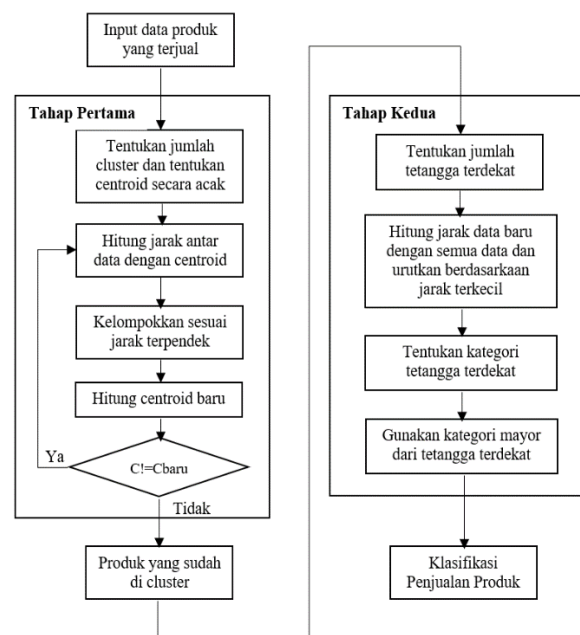
Penelitian tentang klasifikasi penjualan dengan menggunakan algoritma *K-NN* juga pernah dilakukan sebelumnya. Penelitian ini dilakukan di toko yang bergerak di bisnis rokok elektrik yaitu Lombok Vape On. Penelitian ini mengklasifikasikan hasil penjualan untuk mendapatkan efektifitas penjualan per bulannya. Hasil penelitian menunjukkan *K-Fold Validation* 6 memiliki akurasi sebesar 86.48% dan AUC sebesar

0.874 [19]. Penelitian lainnya tentang klasifikasi penjualan produk adalah tentang klasifikasi penjualan produk paket kuota internet. Penelitian ini dilakukan di PT. Akses Lintas Nusantara yang bergerak pada bidang penjualan kuota internet. Penelitian ini menggunakan algoritma *K-NN* untuk melakukan klasifikasi penjualannya. Hasil penelitian ini menunjukkan tingkat akurasi yang mencapai 71.43% [20].

Metode *hybrid data mining* yang diusulkan pada penelitian ini dapat menghasilkan klasifikasi yang memiliki tingkat akurasi yang lebih tinggi dari penelitian yang sudah ada sebelumnya. Dengan melakukan kombinasi antara klusterisasi dan klasifikasi, maka penelitian ini menghasilkan tingkat validitas yang tinggi antara variabel input dan variabel target.

2. Metodologi Penelitian

Penelitian ini mengusulkan kombinasi antara metode klusterisasi dengan metode klasifikasi. Metode klusterisasi digunakan untuk mendapatkan beberapa kelompok produk. Proses klusterisasi dilakukan dengan menggunakan algoritma *K-Means*. Setelah kelompok produk diperoleh, dilanjutkan dengan penggunaan metode klasifikasi untuk mendapatkan klasifikasi penjualan. Proses klasifikasi dilakukan dengan menggunakan algoritma *K-Nearest Neighbors*. Metodologi penelitian yang diusulkan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

Berdasarkan metodologi penelitian pada Gambar 1, dapat dilihat bahwa penelitian menggunakan data produk yang terjual sebagai data input. Penelitian ini menggunakan metode *hybrid data mining*, sehingga terdapat penggabungan metode yang terdiri dari dua tahap pengerjaan, dimana hasil dari tahap pertama akan menjadi input pada tahap kedua.

A. Tahap Pertama

Pada tahap pertama dalam penelitian ini, dilakukan proses klasterisasi dengan menggunakan algoritma *K-Means*. Data produk terjual yang sudah diinputkan, proses klasterisasi dimulai dengan menentukan jumlah *cluster* yang akan dibentuk, dimana data produk akan dikelompokkan menjadi 2 *cluster* yaitu “Banyak Peminat” dan “Sedikit Peminat”. Kemudian dilakukan pembangkitan titik pusat *cluster* atau *centroid* secara acak. Hitung jarak masing-masing data dengan masing-masing *centroid* dengan menggunakan rumus korelasi antar dua objek yaitu *Euclidean Distance* seperti yang ditunjukkan pada Persamaan (1).

$$D_{11} = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} \quad (1)$$

Kemudian data dikelompokkan berdasarkan jarak terpendek antar data dengan *centroid*. Selanjutnya cari *centroid* baru berdasarkan kelompok data yang sudah diperoleh sebelumnya. Jika *centroid* baru tidak sama dengan *centroid* lama, maka dilakukan kembali perhitungan jarak antar data dan *centroid* baru, namun jika *centroid* baru sama dengan *centroid* lama, maka tidak terjadi perubahan anggota kelompok, sehingga *cluster* data produk dapat dihasilkan.

B. Tahap Kedua

Pada tahap kedua dalam penelitian ini, dilakukan proses klasifikasi dengan menggunakan algoritma *K-Nearest Neighbors*. Data produk yang sudah di cluster dari hasil tahap pertama dijadikan sebagai data input untuk tahap kedua. Data yang digunakan adalah data hasil cluster “Banyak Peminat”. Proses klasifikasi dimulai dengan menentukan jumlah tetangga terdekat yaitu 3 tetangga terdekat, karena keputusan yang akan diambil ada 2 keputusan, yaitu “Penjualan Tinggi” dan “Penjualan Rendah”. Kemudian dilakukan perhitungan jarak antara data baru dengan masing-masing data dengan menggunakan rumus *Euclidean Distance* pada persamaan (1). Dari jarak yang diperoleh, data diurutkan berdasarkan jarak terdekat. Selanjutnya ditentukan kategori tetangga terdekat dan gunakan kategori mayor dari tetangga tersebut sehingga klasifikasi penjualan dapat diperoleh dari produk yang banyak peminat.

Dalam penelitian ini, digunakan data produk yang terjual dari satu minimarket yang ada di Kota Padang. Penelitian ini menggunakan 20 data produk yang berhubungan dengan perlengkapan mandi dan mencuci. Data ini digunakan untuk mendapatkan cluster produk dengan menggunakan 3 variabel indikator yaitu stok awal, jumlah terjual dan harga produk. Data produk yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Data Produk sebagai Data Pelatihan

Produk	Stok (unit)	Terjual (unit)	Harga (Rp.)
Shinzui Soap Sakura	216	98	3300
Total Harum Bunga 78 gr	129	129	13400

Dettol Soap Original	285	247	13300
Ciptadent Extra Fresh 130 gr	181	122	5600
Rapika Biang Isi 4	330	66	7300
Nuvo Family 80 gr	270	154	2000
Daia Putih Powder 1 kg	354	83	7000
Sunsilk Shampo Black 170 ml	395	63	17200
Dove Shampo Care 80 ml	153	116	10700
Pepsodent Standar 25 gr	203	203	3400
Ponds AM Cream Day 10 gr	109	24	27900
Mama Lemon 800 ml	205	38	9700
Ekonomi Lemon 900	197	41	1800
Soklin Pemutih 500 gr	109	38	27000
Soklin Antibact Pouch 800 ml	247	67	14900
Rinso Liquid Pouch 800 ml	132	1	23000
Sunsilk Shampo Hair Fall 170 ml	138	1	18400
Clear Shampo Ice Mentol 170 ml	108	28	27100
Rinso Anti Noda 900 gr	336	31	16500
Close Up Green 160 gr	240	37	13200

Dari Tabel 1, dapat dilihat terdapat sebanyak 20 data produk perlengkapan mandi dan mencuci. Data ini adalah data produk yang sudah terjual. Untuk proses perhitungan, langkah pertama yang harus dilakukan adalah memberikan bobot untuk setiap data agar cluster yang dihasilkan lebih tepat. Pemberian bobot yang diberikan memiliki rentang nilai 0 sampai dengan 1. Bobot yang diberikan dapat dilihat pada Tabel 2.

Tabel 2. Bobot untuk Semua Atribut

Variabel Indikator	Atribut	Bobot
Stok Awal	≥ 300 unit	1,0
	≥ 200 unit	0,7
	< 200 unit	0,4
Jumlah Terjual	≥ 160 unit	1,0
	≥ 100 unit	0,7
	< 100 unit	0,4
Harga Produk	≥ Rp. 20.000	1,0
	≥ Rp. 10.000	0,7
	< Rp. 10.000	0,4
Keputusan Penjualan	Penjualan Tinggi	1,0
	Penjualan Rendah	0,7

3. Hasil dan Pembahasan

Setelah menentukan bobot untuk masing-masing atribut, maka data yang ada pada Tabel 1 dirubah menjadi nilai numerik berdasarkan bobot yang ada pada Tabel 2 untuk dapat memperoleh nilai jarak dari masing-masing data. Hasil konversi bobot dapat dilihat pada Tabel 3.

Tabel 3. Data Pelatihan Setelah Diubah Menjadi Numerik

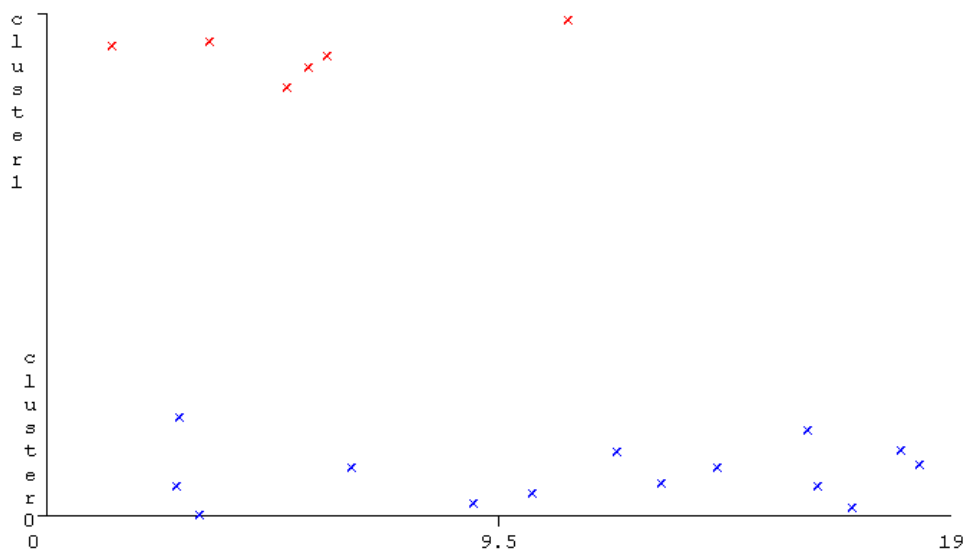
Produk	Stok (unit)	Terjual (unit)	Harga (Rp.)
Shinzui Soap Sakura	0,7	0,4	0,4
Total Harum Bunga 78 gr	0,4	0,7	0,7
Dettol Soap Original	0,7	1,0	0,7
Ciptadent Extra Fresh 130 gr	0,4	0,7	0,4
Rapika Biang Isi 4	1,0	0,4	0,4
Nuvo Family 80 gr	0,7	0,7	0,4
Daia Putih Powder 1 kg	1,0	0,4	0,4

Sunsilk Shampo Black 170 ml	1,0	0,4	0,7
Dove Shampo Care 80 ml	0,4	0,7	0,7
Pepsodent Standar 25 gr	0,7	1,0	0,4
Ponds AM Cream Day 10 gr	0,4	0,4	1,0
Mama Lemon 800 ml	0,7	0,4	0,4
Ekonomi Lemon 900	0,4	0,4	0,4
Soklin Pemutih 500 gr	0,4	0,4	1,0
Soklin Antibact Pouch 800 ml	0,7	0,4	0,7
Rinso Liquid Pouch 800 ml	0,4	0,4	1,0
Sunsilk Shampo Hair Fall 170 ml	0,4	0,4	0,7
Clear Shampo Ice Mentol 170 ml	0,4	0,4	1,0
Rinso Anti Noda 900 gr	1,0	0,4	0,7
Close Up Green 160 gr	0,7	0,4	0,7

Pada tahap pertama, penelitian ini melakukan proses klasterisasi dengan menggunakan algoritma *K-Means* dari data produk yang terjual dari sebuah minimarket. Langkah pertama adalah menentukan jumlah *cluster*

yang akan dibentuk, dimana jumlah *cluster* yang akan dibentuk pada tahap ini sebanyak 2 *cluster* dari 20 data produk. *Cluster* yang ingin dibentuk adalah “Banyak Peminat” dan “Sedikit Peminat”. *Centroid* dibangkitkan secara random untuk masing-masing *cluster* yang ingin dibentuk. Kemudian dilakukan perhitungan jarak antara masing-masing data dengan masing-masing *centroid*.

Data diatas diuji dengan menggunakan aplikasi analisa data yaitu Weka versi 3.8.3. Hasil perhitungan dan pengujian menggunakan Weka, dari 20 data produk terdapat 14 produk (70% dari data) masuk ke dalam “*cluster0*” dan 6 produk (30% dari data) masuk ke dalam “*cluster1*”. *Cluster0* mengacu pada produk yang “Banyak Peminat” dan *cluster1* mengacu pada produk yang “Sedikit Peminat”. Hasil klasterisasi dapat divisualisasikan dalam bentuk *scatter plot* yang dapat dilihat pada Gambar 2.



Gambar2. Scatter Plot Hasil Klasterisasi

Pada tahap kedua, penelitian ini melakukan klasifikasi dengan menggunakan algoritma *K-Nearest Neighbors (K-NN)*. Data yang digunakan pada tahap ini adalah data hasil klasterisasi yang diperoleh dari tahap pertama yaitu data *cluster* “Banyak Peminat”. Dalam *K-NN*, langkah pertama harus ditentukan terlebih dahulu jumlah tetangga terdekat, dimana dalam penelitian ini jumlah tetangga terdekat yang ditentukan adalah 3, karena klasifikasi keputusan yang akan diambil adalah

2, “Penjualan Tinggi” dan “Penjualan Rendah”. Kemudian dilakukan proses perhitungan jarak antara data yang baru dengan semua data pelatihan dan data diurutkan berdasarkan jarak terpendek. Selanjutnya ditentukan kategori tetangga terdekat dan gunakan kategori mayor dari tetangga tersebut. Data pelatihan yang digunakan pada proses *K-NN* dapat dilihat pada Tabel 4.

Tabel 4.Data Pelatihan dari Cluster “Banyak Peminat”

Produk	Stok (unit)	Terjual (unit)	Harga (Rp.)	Keputusan Penjualan
Total Harum Bunga 78 gr	0,4	0,7	0,7	Penjualan Tinggi
Dettol Soap Original	0,7	1,0	0,7	Penjualan Tinggi
Ciptadent Extra Fresh 130 gr	0,4	0,7	0,4	Penjualan Tinggi
Sunsilk Shampo Black 170 ml	1,0	0,4	0,7	Penjualan Rendah
Dove Shampo Care 80 ml	0,4	0,7	0,7	Penjualan Tinggi
Ponds Am Cream Day 10 gr	0,4	0,4	1,0	Penjualan Tinggi
Ekonomi Lemon 900	0,4	0,4	0,4	Penjualan Tinggi
Soklin Pemutih 500 gr	0,4	0,4	1,0	Penjualan Tinggi
Soklin Antibact Pouch 800 ml	0,7	0,4	0,7	Penjualan Rendah

Rinso Liquid Pouch 800 ml	0,4	0,4	1,0	Penjualan Tinggi
Sunsilk Shampo Hair Fall 170 ml	0,4	0,4	0,7	Penjualan Tinggi
Clear Shampo Ic Mentol 170 ml	0,4	0,4	1,0	Penjualan Tinggi
Rinso Anti Noda 900 gr	1,0	0,4	0,7	Penjualan Rendah
Close Up Green 160 gr	0,7	0,4	0,7	Penjualan Rendah

Dari Tabel 4, dapat dilihat bahwa terdapat 2 klasifikasi keputusan penjualan, “Penjualan Tinggi” memiliki 10 produk dan “Penjualan Rendah” memiliki 4 produk. Dalam algoritma K-NN, proses klasifikasi digunakan pada data baru sebagai data pengujian, dimana data tersebut belum memiliki keputusan. Data baru sebagai data pengujian dapat dilihat pada Tabel 5.

Tabel 5. Data Pengujian

Produk	Stok (unit)	Terjual (unit)	Harga (Rp.)
Gatsby Body Wash Refill 250 ml	0,7	0,7	0,7

Berdasarkan data pengujian pada Tabel 5, maka dilakukan pengujian K-NN dengan menggunakan aplikasi Weka 3.8.3. Data pelatihan dan data pengujian dimasukkan secara bersamaan untuk menghasilkan klasifikasi dari data pengujian. Dari proses pengujian yang dilakukan, maka diperoleh produk baru Gatsby Body Wash Refill 250 ml diklasifikasikan sebagai produk “Penjualan Tinggi” dengan *prediction margin* 0.192 seperti yang terlihat pada Gambar 3.

```

Plot : weka.classifiers.lazy.IBk (data_kNN)
Instance: 15
                Produk : Gatsby Body Wash Refill 250 ml
                Stok : 0,7
                Terjual : 0,7
                Harga : 0,7
                prediction margin : 0.1917808219178082
predicted Keputusan Penjualan : Penjualan Tinggi
Keputusan Penjualan : Missing
    
```

Gambar 3. Hasil Klasifikasi Data Pengujian

Dari hasil pengujian, diperoleh hasil ketepatan dalam klasifikasi dari penjualan yaitu mencapai 85.7143 %. Hal ini menunjukkan bahwa penggunaan metode *hybrid data mining* dapat menghasilkan tingkat ketepatan dalam klasifikasi yang tinggi. Sehingga metode *hybrid data mining* yang digunakan sudah bisa menjadi salah satu alternatif solusi dalam melakukan klasifikasi data.

4. Kesimpulan

Dalam penelitian ini, ditemukan masalah yang terjadi pada minimarket khususnya pada bagian penjualan. Sehingga digunakan metode *hybrid data mining* dengan menggabungkan proses klasterisasi dan klasifikasi. Klasterisasi digunakan untuk melakukan pengelompokan produk yang terjual, sementara klasifikasi digunakan untuk mengklasifikasikan penjualan. Penggunaan metode *hybrid* antara algoritma *K-Means* dan *K-Nearest Neighbors* memberikan pengaruh yang signifikan dalam menentukan klasifikasi. Dari hasil pengujian, ketepatan dalam klasifikasi mencapai 85.7143 %. Untuk penelitian selanjutnya, disarankan adanya proses prediksi dari hasil klasifikasi yang sudah ditemukan, sehingga lebih memiliki pengaruh yang besar dalam penggunaan metode *hybrid data mining* ini.

Daftar Rujukan

- [1] B. Hidayah and C. Amin, “Analisis Pola Spasial dan Faktor Pemilihan Lokasi Minimarket di Kabupaten Klaten,” *Media Komun. Geogr.*, vol. 22, no. 2, p. 171, 2021, doi: 10.23887/mkg.v22i2.36806.
- [2] A. Mardiyah and F. A. Safrin, “Persaingan Usaha Warung Tradisional dengan Toko Modern,” *J. Bisnis dan Manaj.*, vol. 8, no. 1, 2021, doi: 10.26905/jbm.v8i1.5454.
- [3] M. Yanto, “Sistem Penunjang Keputusan Dengan Menggunakan Metode Ahp Dalam Seleksi Produk,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 3, no. 1, pp. 167–174, 2021, doi: 10.47233/jteksis.v3i1.161.
- [4] H. Prastiwi, J. Pricilia, and E. Raswir, “Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering,” *J. Inform. Dan Rekayasa Komput.*, vol. 1, no. April, pp. 141–148, 2022.
- [5] E. P. W. Mandala and D. E. Putri, “Prediksi Tingkat Penjualan Sepeda Motor dengan Metode Rough Set,” *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 896, 2021, doi: 10.30865/mib.v5i3.3057.
- [6] F. K. Alvisan, “Clustering Minimarket Untuk Menentukan Jumlah Kebutuhan Pembelian Menggunakan Metode K-Means,” *J. NOE*, vol. 4, no. 2, pp. 160–168, 2021, doi: 10.29407/noe.v4i2.16784.
- [7] E. P. W. Mandala, E. Rianti, and S. Defit, “Classification of Customer Loans Using Hybrid Data Mining,” *JUITA J. Inform.*, vol. 10, no. 1, pp. 45–52, 2022, doi: 10.30595/juita.v10i1.12521.

- [8] M. Abedini, F. Ahmadzadeh, and R. Noorossana, "Customer credit scoring using a hybrid data mining approach," *Kybernetes*, vol. 45, no. 10, pp. 1576–1588, 2016, doi: 10.1108/K-09-2015-0228.
- [9] M. H. Fakhriza and K. Umam, "Analisis Produk Terlaris Menggunakan Metode K-Means Clustering pada PT. Sukanda Djaya," *JIKA (Jurnal Inform.)*, vol. 5, no. 1, pp. 8–15, 2021, doi: 10.31328/jointecs.v6i3.2693.
- [10] H. Syahputra, L. Mayola, and D. Guswandi, "Clustering Tingkat Penjualan Menu (Food and Beverage) Menggunakan Algoritma K-Means," *J. KomtekInfo*, vol. 9, pp. 29–33, 2022, doi: 10.35134/komtekinfo.v9i1.274.
- [11] N. Ananda and R. A. Aras, "Clustering Pengeluaran Tahunan Berbagai Macam Produk Menggunakan Metode K-Means," in *Seminar Nasional Sains dan Teknologi Informasi SENSASI 2021*, 2021, pp. 143–147.
- [12] D. A. Kurnia and A. R. Rinaldi, "Clustering Data Penjualan Produk Makanan pada Toko Toserba Yogya Siliwangi dengan Menggunakan Metode K-Means," *MEANS (Media Inf. Anal. dan Sist.)*, vol. 7, no. 1, pp. 77–84, 2022.
- [13] A. I. Shafarindu, E. Patimah, Y. M. Siahaan, A. W. Wardhana, B. V. Haekal, and D. S. Prasvita, "Klasifikasi Data Penjualan pada Supermarket dengan Metode Decision Tree," *Senamika*, vol. 2, no. 1, pp. 660–667, 2021, [Online]. Available: <https://conference.upnvj.ac.id/index.php/senamika/article/view/1389>.
- [14] F. Maulana, M. Orisa, and H. Zulfia Zahro', "Klasifikasi Data Produk Mebel Aneka Jaya Menggunakan Metode K-Nearest Neighbor Berbasis Web," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 2, pp. 460–466, 2021, doi: 10.36040/jati.v5i2.3782.
- [15] S. P. Dewi and E. R. Nurwati, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, p. 639–648, 2020, doi: 10.47065/bits.v3i4.1408.
- [16] A. Alfani W.P.R., F. Rozi, and F. Sukmana, "Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jipi.v6i1.1910.
- [17] I. S. M. Negara, Purwono, and I. A. Ashari, "Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 dengan Algoritma K-means," *JOINTECS*, vol. 6, no. 3, pp. 153–160, 2021, doi: 10.31000/jika.v5i1.3236.
- [18] I. Nawangsih, R. Puspita, and Suherman, "Implementasi Algoritma K-Means Dalam Mengkategorikan Produk Terlaris Dan Kurang Laris Pada Toko Alfamart Cikarang," *Pelita Teknol.*, vol. 16, no. 1, pp. 79–87, 2021.
- [19] Y. Yahya and W. Puspita Hidayanti, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada 'Lombok Vape On,'" *Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 104–114, 2020, doi: 10.29408/jit.v3i2.2279.
- [20] D. Handoko, H. S. Tambunan, and J. T. Hardinata, "Analisis Penjualan Produk Paket Kuota Internet Dengan Metode K-Nearest Neighbor," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 6, no. 1, p. 111, 2021, doi: 10.30645/jurasik.v6i1.275.