

Implementasi Data Mining Pada Hasil Penjualan Makanan Beku Menggunakan K Means Clustering

Implementation of Data Mining on Froozen Food Sales Results Using K Means Clustering

Nenna Irsa Syahputri¹, Rismayanti², Siti Sundari^{3*}

^{1,2,3}Prodi Teknik Informatika, Universitas Harapan Medan

E-mail: ¹nenna.ziadzha@gmail.com, ²risma.stth@gmail.com, ³sundaristth@gmail.com

Abstrak

Menjaga stok persediaan barang agar tidak ada barang yang kosong termasuk salah satu cara untuk menjaga kepuasan pelanggan. Dalam persaingan dunia bisnis pada saat ini, kita dituntut untuk selalu mengembangkan bisnis agar bertahan dalam persaingan, khususnya dalam persaingan penjualan menuntut para pengusaha untuk menemukan suatu pola yang dapat meningkatkan penjualan dan pemasaran di perusahaan, salah satunya adalah dengan pemanfaatan data penjualan. Menerapkan teknik data mining klastering sehingga dapat membantu toko NCekma Frozen dalam menentukan strategi penentuan stok makanan beku menggunakan algoritma K Means.

Kata kunci: Data mining, algoritma K Means

Abstract

Maintaining inventory stock so that there are no empty items is one way to maintain customer satisfaction. In today's competitive business world, we are required to always develop our business in order to survive in the competition, especially in sales competition, it requires entrepreneurs to find a pattern that can increase sales and marketing within the company, one of which is by utilizing sales data. Applying clustering data mining techniques so that it can help NCekma Frozen stores in determining strategies for determining frozen food stocks using the K Means algorithm.

Keywords: Data mining, K Means algorithm.

1. PENDAHULUAN

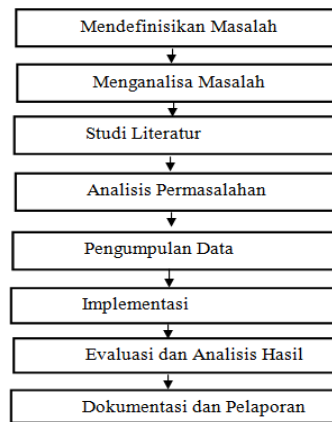
Dalam persaingan dunia bisnis pada saat ini, kita dituntut untuk selalu mengembangkan bisnis agar bertahan dalam persaingan, khususnya dalam persaingan penjualan menuntut para pengusaha untuk menemukan suatu pola yang dapat meningkatkan penjualan dan pemasaran di perusahaan, salah satunya adalah dengan pemanfaatan data penjualan.”[1]. Beberapa permasalahan pun timbul seperti menumpuknya produk yang tidak laku terjual di gudang perusahaan. Hal ini mengakibatkan tidak optimalnya laba yang didapat dari hasil penjualan produk. Penyebabnya karena kurang tepatnya keputusan yang diambil pihak manajemen terkait dalam hal menentukan strategi terhadap persediaan produk dan cara memasarkannya. Untuk mengatasi masalah tersebut perusahaan

harus menganalisa secara tepat dengan didukung informasi yang cukup banyak untuk mengambil kesimpulan guna sebagai bahan pertimbangan dalam pengambilan keputusan. Ketersediaan data yang cukup besar tidak dimanfaatkan oleh perusahaan secara optimal untuk mendapatkan informasi tersembunyi terkait untuk pengembangan perusahaan. Belum adanya Sistem pengambilan keputusan dan metode yang digunakan untuk strategi bisnis dalam mengoptimalkan laba penjualan.

Data mining adalah analisa terhadap suatu data untuk menentukan hubungan yang jelas serta menyimpulkan yang belum diketahui sebelumnya dengan cara terkini dan berguna bagi pemilik data tersebut [2]. Metode k-means membagi data menjadi beberapa kelompok sehingga data dengan karakteristik yang sama berada pada cluster yang sama dan data dengan karakteristik yang berbeda berada pada cluster yang berbeda [3]. Dari berbagai permasalahan diatas, penulis akan mengimplementasikan konsep data mining menggunakan Algoritma K-Means untuk mengkluster data barang untuk strategi penentuan stok makanan beku. Data mining merupakan penggalian data yang tersembunyi dari database. Proses Clustering merupakan suatu proses pengelompokan berdasarkan atas prinsip kesamaan kelas serta mengurangi kesamaan antar kelas. Berbagai algoritma dalam clustering telah dikembangkan untuk menghasilkan kinerja yang baik. Keakuratan perkiraan penjualan memiliki dampak yang besar pada penjualan. Peramalan penjualan menjadi factor penting pada manajemen dimulai dari pengecer sampai pada distributor, produsen dan supplier. Hasil perkiraan penjualan menggunakan metode k-means yang tepat dan akurat dapat menjadi jembatan antara banyak penawaran dan permintaan sehingga mampu mengurangi biaya dan mempertahankan jumlah stok barang [4][5][6].

Cabang ilmu komputer yang relatif baru adalah Data Mining. Masih menjadi perdebatan dalam menempatkan data mining pada bidang ilmu yang mana. Ada beberapa pihak yang menyatakan data mining tidak lebih dari analisis statistik yang berjalan diatas database dan ada pihak yang mempunyai pendapat bahwa database berperan penting dalam data mining terutama dalam optimasi querynya[7]. Clustering merupakan suatu kelompok entitas yang memiliki persamaan dan memiliki perbedaan entitas dari kelompok lain. Cara kerja teknik kluster yaitu dengan mengelompokkan objek-objek data (entitas, kejadian, pola, unit, hasil observasi) atau melakukan pemisahan/ segmentasi/ pemecahan data kedalam sejumlah kluster tertentu. Teknik 23 kluster dapat di aplikasikan keberbagai bidang ilmu di antaranya : Teknik, Ilmu Komputer, Medis, Astronomi, Sosial, dan Ekonomi [8].

2. METODOLOGI PENELITIAN



Gambar 2.1 Kerangka Kerja

Dalam penelitian penulis menggunakan beberapa metode penelitian untuk mendapatkan data-data yang diperlukan. Beberapa metode yang digunakan adalah sebagai berikut :

1. Studi Literatur

Studi literatur yaitu metode pengumpulan data dengan mencari referensi dari buku-buku dan media internet yang berkaitan dengan penelitian. Adapun bahan referensi adalah datamining, Metode K Means.

2. Analisis Permasalahan

Menganalisis masalah membuat sistem kluster pengelompokan transaksi barang menggunakan algoritma K-means pada toko Ncekma Dimsum dalam menentukan strategi penentuan stok barang..

3. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data ditoko Ncekma Dimsum.

4. Implementasi

Merupakan tahap mengimplementasikan metode K Means dan Aplikasi Rapid Miner dalam mengkluster data barang.

5. Evaluasi dan Analisis Hasil

Pada tahap ini dilakukan evaluasi dan analisis hasil yang terdapat melalui implementasi Metode K Means

6. Dokumentasi dan Pelaporan

Pada tahap ini dilakukan dokumentasi dan penyusunan laporan hasil evaluasi dan analisis serta implementasi metode

3. HASIL DAN PEMBAHASAN

3.1. Hasil dan Pembahasan

Penelitian dilakukan di Ncekma Frozen di Desa Kolam, Lamanya waktu yang dibutuhkan untuk menyelesaikan penelitian ini selama 2 bulan mulai Agustus sampai September 2021. Data yang digunakan dalam penelitian adalah data

Penjualan Frozen Food dari bulan Agustus 2021. Kumpulan data yang diperoleh penulis digunakan sebagai data masukan dalam membuat model aturan menggunakan algoritma K-Means dengan 2 cluster dan menggunakan software Rapidminer. Analisis data yang dilakukan untuk mengolah data penjualan Frozen Food dengan memperoleh data penjualan dari bulan Agustus 2021. Dibagi menjadi 2 bagian yaitu cluster sanga tlaris dan laris. Sehingga karakteristik data tersebut dapat dipahami dan bermanfaat untuk solusi permasalahan dari penelitian.

3.2. Analisis Perhitungan Data Uji

Data penjualan Frozen Food di bulan Agustus terdapat pada tabel berikut

Tabel 4.1 Tabel Penjualan Frozen Food di Bulan Agustus 2021

No	Nama Barang	Stok Awal	Total Penjualan	Stok Akhir
1	Sosis Okey	50	48	2
2	Sosis Asimo	45	20	25
3	Sosis Belfood	21	12	9
4	Sosis Bonbon	20	14	6
5	Sosis Vigo	20	10	10
6	Naget Okey	45	40	5
7	Naget Asimo	25	20	5
8	Naget Belfood	25	15	10
9	Kentang	20	14	6
10	Minipau	15	3	12
11	Bakso Champ	30	22	8
12	Burger Champ	35	28	7
13	Bakso Amanah	40	35	5
14	Nuget Amanah	25	20	5
15	Tempura Asyifa	45	37	8
16	Otak-Otak Asyifa	40	33	7
17	Taiso Rayhan	30	24	6
18	Bakso Rayhan	30	22	8
19	Bakso Bunga Barokah	40	34	6
20	Naget Kaki Naga	30	13	17
21	Lumpia Maya	20	13	7
22	Kulit Pangsit	35	25	10

3.3 Perhitungan K Means

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

1. Tentukan nilai k nya sebagai jumlah klaster 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 :

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

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.

b. Iterasi 1

Menentukan pusat centroid awal Centroid akan dibagi menjadi 2 kelompok penentuan centroid awal dilakukan dengan mengacak (random) data yang sudah ada. Berikut ini merupakan Centroid awalnya.

Tabel 3.1 Centroid awal

Keterangan	Stok Awal	Stok Terjual	Stok Akhir
C1	50	48	2
C2	30	22	8

Data yang dipilih untuk centroid awal adalah data ke 1 dan data ke 11.

1. Perhitungan jarak dengan pusat cluster Perhitungan jarak dari data ke-i terhadap pusat centroid adalah :

$$D(1,1) = \sqrt{(50 - 50)^2 + (48 - 48)^2} = 0$$

$$D(1,2) = \sqrt{(50 - 30)^2 + (48 - 22)^2} = 28,07134$$

,

,

,

$$(22,2) = \sqrt{(35 - 15)^2 + (25 - 3)^2} = 24,35957$$

Tabel 3.2 Pengelompokan data ke-1

No	C1	C2
1	1	
2		1
3		1
4		1
5		1
6	1	

7		1
8		1
9		1
10		1
11		1
12		1
13	1	
14		1
15	1	
16		1
17		1
18		1
19	1	
20		1
21		1
22		1

Menghitung nilai WCV (Within Cluster Variation) Rumus menghitung nilai WCV yaitu dengan memangkatkan hasil perhitungan jarak dan menjumlahkan seluruh hasil perpangkatan dari hasil perhitungan jarak. $WCV = 0 + 926 + 1995 + \dots + 593 = 24.901$

c. Iterasi 2

Iterasi/perulangan ke-2 merupakan proses penghitungan kembali dikarenakan masih ada data yang berpindah cluster.

Tabel 3.2 Centroid Baru

Keterangan	Stok Awal	Stok Terjual	Stok Akhir
C1	44	38,8	5,2
C2	28,06666667	18,867	9,2

Tabel 3.2 adalah centroid baru yang dihasilkan dari rata-rata dari 2 cluster.

1. Perhitungan jarak dengan pusat cluster Perhitungan jarak dari data ke-i terhadap pusat centroid adalah :

$$D(1,1) = \sqrt{(50 - 44)^2 + (48 - 38,8)^2} = 11,44028$$

$$D(1,2) = \sqrt{(50 - 28,066)^2 + (48 - 18,867)^2} = 26,20765$$

$$D(22,2) = \sqrt{(35 - 15)^2 + (25 - 3)^2} = 21,55016$$

Tabel 3.3 Hasil Perhitungan Jarak Data terhadap Cluster Iterasi 2

No	Nama Barang	SA	TP	SAT	C1	C2	Jarak Min	Cluster	MD ²
1	Sosis Okey	50	48	2	11,44028	26,20765	11,44027972	c1	130,88
2	Sosis Asimo	45	20	25	27,321786	20,76266	20,76265935	c2	431

3	Sosis Belfood	21	12	9	35,520135	32,37651	32,37651004	c2	1048
4	Sosis Bonbon	20	14	6	34,520718	31,69481	31,69481035	c2	1005
5	Sosis Vigo	20	10	10	37,795238	34,96727	34,96726822	c2	1223
6	Naget Okey	45	40	5	1,5748016	19,82	1,574801575	c1	2
7	Naget Asimo	25	20	5	26,729759	23,76232	23,76232302	c2	565
8	Naget Belfood	25	15	10	30,829856	26,78071	26,78070595	c2	717
9	Kentang	20	14	6	34,520718	31,69481	31,69481035	c2	1005
10	Minipau	15	3	12	46,571236	45,5203	45,52030482	c2	2072
11	Bakso Champ	30	22	8	22,047222	17,88687	17,88686678	c2	320
12	Burger Champ	35	28	7	14,173214	12,86682	12,86682095	c2	166
13	Bakso Amanah	40	35	5	5,5208695	15,93397	5,520869497	c1	30
14	Nuget Amanah	25	20	5	26,729759	23,76232	23,76232302	c2	565
15	Tempura Asyifa	45	37	8	3,4756294	15,18741	3,475629439	c1	12
16	Otak-Otak Asyifa	40	33	7	7,2718636	12,99512	7,271863585	c2	53
17	Taiso Rayhan	30	24	6	20,388232	17,52899	17,52898805	c2	307
18	Bakso Rayhan	30	22	8	22,047222	17,88687	17,88686678	c2	320
19	Bakso Bunga Barokah	40	34	6	6,2992063	14,46272	6,299206299	c1	40
20	Naget Kaki Naga	30	13	17	43,011626	37,06366	37,06366097	c2	1374
21	Lumpia Maya	20	13	7	46,368092	41,82451	41,82450622	c2	1749
22	Kulit Pangsit	35	25	10	28,600699	21,55016	21,55015905	c2	464

Tabel 3.4 Pengelompokan data ke-2

No	C1	C2
1	1	
2		1
3		1
4		1
5		1
6	1	
7		1
8		1

9		1
10		1
11		1
12		1
13	1	
14		1
15	1	
16		1
17		1
18		1
19	1	
20		1
21		1
22		1

Centroid baru dihitung menggunakan rumus seperti rumus iterasi per-1 dan karena tidak ada data yang berpindah cluster dan cluster ke- 1 dan ke-2 hasilnya sama, maka proses perhitungan centroid yang baru dihentikan dan berakhir pada iterasi ke-2

4. KESIMPULAN

Kesimpulan yang dapat di ambil diantaranya

1. Implementasi K-Means Clustering berdasarkan data uji dapat mengelompokkan data produk sangat laku, laku. Contoh produk yang sangat laku ditunjukkan pada penjualan Sosis Okey, Naget Okay, Bakso Amanah, Tempura Asyifa, Bakso Bunga Barokah. Contoh produk laku ditunjukkan pada penjualan Sosis Asimo, Sosis Belfood, Sosis Bonbon, Sosis Vigo, Naget Asimo, Naget Belfood, Kentang, Minipau, Bakso Champ, Burger Champ, Naget Amanah, Otak2 AsyifaTaiso Rayhan, Bakso Rayhan, Naget Kaki Naga, Lumpia Maya, Kulit Pangsit.
2. Menggunakan algoritma K-Means Cluster data dapat di kelompokkan kedalam 2 kelas cluster dari 22 data penjualan yang sudah ditambahkan.

DAFTAR PUSTAKA

- [1] Siregar, M. H. (2018). Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan). Jurnal Teknologi Dan Open Source, 1(2), 83-91. <https://doi.org/10.36378/jtos.v1i2.24>.
- [2] Metisen, B. M., & Sari, H. L. (2015). Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila. J. Media Infotama, vol. 11, pp. 110-118.

- [3] Rohmawati, N., Defiyanti, S., & Jajuli, M. (2015). Implementasi Algoritma K-MEANS dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *Jurnal Ilmiah Teknologi Informasi Terapan*, 1(2), 62-67.
- [4] M. Sanwlan, "FORECASTING SALES THROUGH TIME SERIES CLUSTERING," vol. 3, no. 1, pp. 39– 56, 2013.
- [5] V. Shrivastava, P. Arya, and M. T. S. Systems, "International Journal of Computing , Communications and Networking Available Online at <http://warse.org/pdfs/ijccn04122012.pdf> A Study of Various Clustering Algorithms on Retail Sales Data," 2012.
- [6] A. Benet-zepf, J. A. Marin-garcia, and I. Küster, "Clustering the mediators between the sales control systems and the sales performance using the AMO model : A narrative systematic literature review," vol. 14, no. 2, pp. 387–408, 2018.
- [7] Fadli, A. (2011). *Konsep Data Mining. Pohon Keputusan*, vol. 1, pp. 1-9.
- [8] Suprayogi. (2013). *Data Mining Clustering*. *J. Chem. Inf. Model.*, vol. 53, pp. 1689-1699.