

IMPLEMENTASI ALGORITMA *K-MEAN* DALAM MENENTUKAN BAHAN YANG LARIS TERJUAL PADA TOKO BANGUNAN UD.HERSUK

Oleh :

Nopi Purnomo¹), Fery Irawan²), Mustopa Husein Lubis³), Akhir Abadi Tanjung⁴)

^{1,2,3}Institut Teknologi dan Sains Paluta

²NamaFakultas, nama Perguruan Tinggi (penulis 2)

nopipurnomo2020@gmail.com

Fery.irawan1922@gmail.com

mustopahusein1302@gmail.com

tanjungakhirabadi@gmail.com

Abstrak

Data merupakan suatu hal yang sangat dibutuhkan dalam teknologi informasi. Tujuan penelitian ini adalah untuk mengelompokkan jenis data produk berdasarkan dari tingkat penjualan dan untuk mengetahui produk mana yang memiliki tingkat pembelian yang rendah, sedang dan tinggi pada Toko Bangunan UD.Hersuk Kota Padang. Metode penelitian yang digunakan ada 5 yaitu penentuan masalah, perumusan masalah, kajian pustaka, pemecahan masalah dan penarikan kesimpulan. Dalam penelitian ini proses clustering secara hirarki menggunakan Algoritma *K-Means*. Proses komputasi yang digunakan untuk mengolah data adalah program Tanagra 1.4.42 for windows. Data di ambil dari Toko Bangunan UD.Hersuk Kota Padang tentang data pembelian barang.

Kata Kunci: Implementasi, Bahan yang Laris Terjual, Algoritma *K-Means*, Tanagra.

1. PENDAHULUAN

UD.Hersuk adalah sebuah toko yang bergerak dibidang penjualan bahan bangunan, UD.Hersuk ini masih menggunakan cara manual untuk menghasilkan laporan keuangan yang diperlukan. Semua transaksi dicatat pada akhir jam kerja tiap harinya sehingga tidak dapat memberikan informasi secara cepat dan akurat.

Oleh karena itu, dalam penelitian ini dirancang dan dibuat sebuah sistem informasi data mining yang membantu permasalahan untuk mendapatkan informasi yang diperlukan secara cepat, akurat dan diharapkan juga dapat menghindari atau meminimalkan terjadinya kesalahan yang timbul akibat pencatatan aktivis toko bangunan, serta menghasilkan laporan yang berguna untuk kemajuan toko bangunan

Perancangan desain sistem berbasis data mining ini menggunakan Algoritma *K-Mean* dengan menggunakan software Tanagra. Dari hasil pengujian program menunjukan bahwa program ini mampu memproses data toko bangunan yang berhubungan dengan pembelian dan penjualan secara otomatis dan menghasilkan informasi berupa laporan manajerial (laporan pembelian, penjualan, dan hutang), dan laporan keuangan (laba rugi dan neraca) yang akurat (sesuai dengan proses manualnya). Untuk itu penelitian ini bermaksud membantu menyelesaikan masalah yang dihadapi oleh Toko Bangunan UD.Hersuk terkait dengan pencatatan penjualan barang agar atasannya bisa menerima data yang

akurat.

Data Mining adalah suatu proses atau kegiatan dalam mengumpulkan berbagai jumlah data yang berukuran besar kemudian mengekstraksi data tersebut menjadi suatu informasi yang sangat penting dan dapat digunakan nantinya [1].

K-Means merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk melakukan pengelompokan atau Clustering suatu data [2].

K-Means pada penelitian terdahulu digunakan sebagai Clustering dengan menggunakan Metode *K-Means* untuk menentukan status gizi balita. Hasil dari penelitian adalah membandingkan hasil pengelompokan algoritma *K-Means* dan tabel *Growth Chart*. Dengan membandingkan hasil pengelompokan menggunakan tabel *Growth Chart* dan algoritma *K-Means* didapat 17 data yang memiliki kelompok yang sama. Dari angka ini dapat disimpulkan bahwa algoritma *K-Means* hanya memiliki nilai akurasi 34% benar [3]. Sedangkan Gustientiedinaa, dkk. (2019) Hasil metode ini adalah dapat mengelompokkan data obat-obatan di Rumah Sakit Umum Daerah Pekanbaru yang dapat digunakan sebagai referensi dalam pengambilan keputusan dalam perencanaan dan pengendalian pasokan medis di rumah sakit tersebut [4].

2. METODE PENELITIAN

Penelitian ini menggunakan Kerangka Kerja

sebagai berikut :



Gambar 1. Kerangka Penelitian

2.1 Tahapan Penelitian

2.1.1 Penelitian Pendahuluan

Penelitian ini berhubungan dengan masalah yang dihadapi oleh toko bangunan UD.Hersuk dalam melakukan pencatatan terhadap penjualan barang pada toko bangunan UD.Hersuk tersebut masih melakukan secara manual, sehingga membuat para pegawai untuk menyampaikan laporan penjualan yang tepat dan akurat kepada atasannya.

2.1.2 Pengumpulan Data

Pengumpulan data yang kami lakukan juga memerlukan proses, agar bisa mendapatkan data yang tepat dan sesuai dengan kebutuhan agar mendapatkan hasil yang maksimal.

a. Waktu Penelitian

Penelitian dilakukan pada bulan April 2021.

b. Tempat Penelitian

Penelitian ini kami lakukan bertempat pada Toko Bangunan UD.Hersuk di Jln. Pisang Pauh kota padang.

2.1.3 Analisis

Berikut ini adalah proses analisa data yang dilakukan :

1. Patern evolution

Hasil data yang dilakukan akan memunculkan informasi yang dapat membentuk pola tertentu. Dalam tahap ini dilakukan proses identifikasi untuk melihat pola data yang menarik dan menganalisis hasil evaluasi.

2. Knowledge presentation

Pola informasi yang dihasilkan dapat menjadi pengetahuan baru yang dapat dipresentasikan kepihak toko bangunan UD.Hersuk.

2.1.4 Perancangan

Tahap perancangan ini dilakukan pengumpulan

data yang bertujuan agar data yang didapat bisa diolah dengan semestinya, agar bisa mendapatkan hasil yang maksimal yang dapat disampaikan kepada atasan.

2.1.5 Implementasi

Data yang telah ada diolah dengan menggunakan algoritma k-means agar dapat mengelompokkan data yang seharusnya, terhadap penjualan barang yang terjadi.

2.1.6 Pengujian

Pengujian dilakukan dengan menggunakan software data mining yaitu tanagra, dan bisa membuktikan apakah hasil manual bisa sama dengan apa yang dihasilkan oleh software.

3. HASIL DAN PEMBAHASAN

Data dalam penelitian ini yang diambil dari Toko Bangunan UD.Hersuk kota padang yaitu data tentang bahan toko bangunan pada bulan Oktober 2021. Data yang digunakan dalam permasalahan ini ada 2 variabel, dimana variabel-variabel tersebut adalah harga barang dan jumlah, sedangkan obyek-obyek yang digunakan 28 jenis produk atau minuman seperti : Paku beron, Les tropille, Besi beton 8mm”12m”, Seng SNI, Triplek 9mm, Kayu balok meranti, Asbes ukuran 300 * 105, Besi beton 12mm”12m”, asbes 12mm”12m”, Triplek 6mm, Kayu galar meranti(5*10), Triplek 4mm, Paralon ¾, Dulux putih, Catylac, Keramik supermilan 25*33, Semen, Gembok prohex, Plastik pagar motif, Seng merah, Impra warna, GRC 4mm, Genteng Surya 1*4 T.0.3, Avitex 5 kg.

Dalam penelitian ini penulis memproses clustering dengan menggunakan metode K-Means atau algoritma K-Means.

3.1 Proses Analisis Cluster

3.1.1 Menentukan Kelompok Data

Jarak tiap objek (jenis bahan) ke masing-masing centroid menggunakan rumus kolerasi antar dua objek yaitu *Euclidean Distance*.

Tabel 1. Data Bahan Bangunan Bulan Oktober 2021

Data Ke-	Nama Barang	M ke-	Harga * 1000 [X]	Jumlah [Y]
1	Paku Beron	M ₁	21.8	20
2	Les tropille	M ₂	18.8	12
3	Besi beton 8mm”12”	M ₃	36	36
4	Seng SNI	M ₄	17.4	3
5	Triplek 9mm	M ₅	100	20
6	Kayu balok meranti	M ₆	92.4	24
7	Asbes ukuran 300*105	M ₇	95.9	24
8	Besi beton 12mm”12”	M ₈	95	20

9	Asbes 210*105	M ₉	61.9	24
10	Triplek 6mm	M ₁₀	83	20
11	Kayu galar meranti	M ₁₁	78.7	30
12	Karawang	M ₁₂	4.1	1
13	Triplek 4mm	M ₁₃	55.2	15
14	Paralon ¾	M ₁₄	45	24
15	Dulux putih	M ₁₅	48	24
16	Catylac	M ₁₆	83.1	10
17	Keramik supermilan 25*33	M ₁₇	55.5	14
18	Semen	M ₁₈	56.7	14
19	Gembok prohex	M ₁₉	30.2	6
20	Plastik pagar motif	M ₂₀	31,7	2
21	Seng merah	M ₂₁	48	60
22	Impra warna	M ₂₂	68.5	20
23	GRC 4mm	M ₂₃	67	40
24	Genteng Surya 1*4 T.0.3	M ₂₄	40	6
25	Avitex 5kg	M ₂₅	78.8	2

M11	58,7	58,7
M12	33,07	33,07
M13	38,26	38,26
M14	25,71	25,71
M15	28,63	28,63
M16	66,19	66,19
M17	38,44	38,44
M18	40,04	40,04
M19	26,08	26,08
M20	30,35	52,41
M21	41,04	41,04
M22	49,52	12,5
M23	48,05	13,45
M24	31,24	43,27
M25	65,13	43,27

Maka kita akan mendapatkan hasil dari masing-masing titik pusat (C), yaitu :

- a. M1,M2,M3,M4,M12,M14,M15,M19,M20,M21,M24 merupakan anggota dari C1
- b. M5,M6,M7,M8,M9,M10,M11,M13,M16,M17,M18,M22,M23,M25 merupakan anggota dari C2

2. Iterasi ke 2

Pada proses intersi ke 2 ini, peneliti harus menentukan posisi centroid yang baru dengan cara menghitung nilai rata-rata dari data-data yang ada pada centroid yang sama, yaitu dengan cara :

3.1.2. Membuat cluster

1. Iterasi ke 1

Pusat awal cluster atau centroid yang ditentukan secara random, untuk penentuan awal cluster diasumsikan :

Pusat Cluster 1 : (20 ; 30)

Pusat Cluster 2 : (76 ; 30)

Sampel data dengan pusat cluster yang dimisalkan dengan M.

Tabel 2. Hasil perhitungan Euclidean distance Iterasi 1

	C1	C2
M1	10,16	55,11
M2	18,04	59,96
M3	17,09	59,96
M4	27,12	64,52
M5	80,62	26,00
M6	72,65	17,46
M7	76,14	20,78
M8	75,66	21,47
M9	42,33	15,32
M10	63,79	63,79

Tabel 3. Hasil perhitungan Euclidean distance Iterasi 2

	C1	C2
M1	9.50	54.75
M2	13.44	58.27
M3	19.03	43.67
M4	19.98	61.48
M5	69.04	23.45
M6	61.73	16.40
M7	65.21	19.80
M8	64.04	18.45
M9	31.55	15.29
M10	52.02	6.45
M11	49.27	10.44
M12	31.63	74.83

M13	24.34	21.88
M14	15.38	31.83
M15	18.15	28.86
M16	52.65	11.78
M17	24.77	21.83
M18	25.96	20.67
M19	11.67	48.45
M20	15.65	48.24
M21	45.64	49.32
M22	37.57	8.05
M23	42.37	22.36
M24	14.71	39.06
M25	50.29	17.92

Maka kita akan mendapatkan hasil dari masing-masing titik pusat (C), yaitu :

- a. M1,M2,M3,M4,M12,M14,M15,M19, M20,M21,M24 merupakan anggota dari C1
- b. M5,M6,M7,M8,M9,M10,M11,M13, M16,M17,M18,M22,M23,M25 merupakan anggota dari C2

Menentukan centroid (titik pusat baru) untuk mencari perbandingan jarak terdekat dengan iterasi selanjutnya. Proses akan berhenti jika perbandingan iterasi 1 dan iterasi 2 sama (pada iterasi 1 anggota C₁ (M₁, M₂, M₃, M₄, M₁₂, M₁₄, M₁₅, M₁₉, M₂₀, M₂₁, M₂₄, dan C₂ (M₅, M₆, M₇, M₈, M₉, M₁₀, M₁₁, M₁₃, M₁₆, M₁₇, M₁₈, M₂₂, M₂₃, M₂₅)) dan pada iterasi 2 C₁ (M₁, M₂, M₃, M₄, M₁₂, M₁₄, M₁₅, M₁₉, M₂₀, M₂₁, M₂₄, dan C₂ (M₅, M₆, M₇, M₈, M₉, M₁₀, M₁₁, M₁₃, M₁₆, M₁₇, M₁₈, M₂₂, M₂₃, M₂₅))

Setelah cluster terbentuk, tahap selanjutnya yaitu memberi nama spesifik untuk menggambarkan isi cluster tersebut. Dari kedua cluster yang terbentuk kita dapat mengklasifikasikan sebagai berikut :

Cluster pertama: Jenis produk(bahan) yang tidak laris, sehingga dapat kita golongkan menjadi kelompok produk dengan tingkat pembelian rendah, anggotanya adalah Paku Beron, Les tropille, Besi beton 8mm”12”, Seng SNI, Karawang, Paralon ¾,v, Gembok prohex,

Plastik pagar motif, Seng merah, Genteng Surya 1*4 T.0.3.

Cluster kedua: Jenis bahan bangunan yang laris, sehingga dapat kita golongkan menjadi kelompok produk dengan tingkat pembelian tinggi, anggotanya adalah Triplek 9 mm, Kayu balok meranti, Asbes ukuran 300*105, Besi beton 12mm”12m”, Asbes 210*105, Triplek 6mm, Kayu galar meranti, Triplek 4 mm, Catylac, Keramik Supermilan 25*33, Semen, Imprap warna, GRC 4 mm, Avitex 5kg.

4. KESIMPULAN

Berdasarkan uraian pada bab-bab sebelumnya, maka penulis dapat mengemukakan beberapa kesimpulan antara lain :

1. Proses cluster secara hirarki dengan menggunakan metode *K-means* mampu menunjukkan keterkaitan setiap jenis barang.
2. Hasil analisis cluster_K-means dari 25 jenis barang yang dapat dikelompokkan menjadi dua cluster_k-means, yaitu:
 - a. Cluster_k-means2, adalah kelompok produk pembelian tinggi yaitu lebih dari separuh jumlah barang.
 - b. Cluster_k-means1, adalah kelompok produk pembelian terendah.

REFERENSI

- [1] Oscario, Jasmir., Novianto, Y. (2019). Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD Sariputra Jambi). *Jurnal STIKOM Dinamika Bangsa*, 14(2), 141-152. <http://doi.org/10.33998/processor.2019.10.637>
- [2] Dinata, R. K., Safwandi, Hasdyna, N., Azizah, N. (2020). Analisis *K-Means Clustering* pada Data Sepeda Motor. *Informatics Journal*. 5(1), 10-17.
- [3] Dhuhita, W. M. P. (2015). *Clustering Menggunakan Metode K-Means* untuk Menentukan Status Gizi Balita. *Jurnal Informatika*. 15(2), 160-174.
- [4] Gustientiedinaa., Adiyaa, M. H., Desnelitab, Y. (2019). Penerapan Algoritma *K-Means* Untuk *Clustering* Data Obat-Obatan Pada RSUD Pekanbaru. *Jurnal Nasional Teknologi dan Sistem Informasi*. 5(1), 17-24. <https://doi.org/10.25077/TEKNOSI.v5i1.2019.17-2>