



Klasifikasi Data Penjualan Menggunakan Algoritma K-Means Dan Analytic Hierarchy Process

Bibit Sudarsono, Umi Faddillah, Ayuni Asistyasari, Yosep Nuryaman

Universitas Bina Sarana Informatika

Jl. Kramat Raya No. 98 Senen. Jakarta Pusat Alamat, Telp/Fax Institusi

bibit.bbs@bsi.ac.id

Abstrak-- Dengan adanya Covid 19 yang telah melanda beberapa tahun terakhir ini membuat penjualan di berbagai bidang mengalami dampak negatif tak terkecuali bagi Swalayan X yang mengalami penurunan di berbagai cabang yang terbagai ke beberapa daerah. Sebagai bahan pertimbangan, mereka melakukan evaluasi terhadap penjualan cabang-cabang tersebut. Dengan mengklasifikasikan data penjualan yang ada diharapkan mereka mampu melihat mana kelompok cabang swalayan x yang kurang baik mana yang sudah baik. Namun perlu adanya teknik klasifikasi yang baik agar evaluasi yang dilakukan berdasarkan hasil perhitungan yang baik. Oleh sebab itu penulis mencoba menggunakan Algoritma K-Mean dan Algoritma AHP untuk mengklasifikasikan data penjualan yang ada. Algoritma K-Mean dan Algoritma AHP merupakan algoritma yang mampu mengklusterisasi sekumpulan data. Dengan mengklusterisasi toko-toko berdasarkan kedekatan hasil penjualan selama 2 tahun terakhir yang naik turun dengan menggunakan 2 algoritma tersebut nantinya akan mampu melihat toko mana tergolong baik dan mana yang tergolong kurang baik. Berdasarkan hasil perbandingan dari hasil perhitungan, didapatkan bahwa hasil terbaik yaitu menggunakan algoritma K-Mean dengan k2 pada literasi ke 3 dengan rasio 0,04926.

Kata kunci: Kmean, AHP, Data mining

Abstract - With the Covid 19 which has hit the last few years, sales in various fields have a negative impact, including Supermarket X, which has experienced a decline in various branches in various regions. For consideration, they evaluate the sales of these branches. By classifying the existing sales data, it is hoped that they will be able to see which self-service branch group x is not doing well, which one is already good. However, it is necessary to have a good classification technique so that evaluations are carried out based on good calculation results. Therefore, the author tries to use the K-Mean Algorithm and the AHP Algorithm to classify the existing sales data. The K-Mean Algorithm and the AHP Algorithm are algorithms that are able to cluster a set of data. By clustering the stores based on the proximity of the sales results for the last 2 years which has gone up and down using these 2 algorithms, we will be able to see which stores are classified as good and which are not. Based on the comparison results from the calculation results, it was found that the best results were using the K-Mean algorithm with k2 in the 3rd literacy with a ratio of 0.04926.

Keywords: K-mean, AHP, Data Mining

I. PENDAHULUAN

Covid 19 membuat penjualan di berbagai bidang mengalami penurunan. Penurunan penjualan tersebut jika tidak diantisipasi dengan baik maka akan berakibat fatal bagi berlangsungnya perusahaan tersebut. Antisipasi yang dapat dilakukan diantaranya dengan memberikan diskon produk, paket penjualan hingga promo-promo lainnya yang mampu menarik pelanggan untuk berbelanja/membeli barang-barang yang ada di toko.

Swalayan x yang berkantor pusat di Jakarta memiliki ribuan cabang di seluruh Indonesia. Dengan promo-promo yang mereka adakan, mereka berharap penjualan di berbagai swalayan yang ada dapat berjalan dengan baik. Namun pada kenyataannya dikarenakan imbas pandemic yang ada, tidak semua swalayan dapat mencapai target penjualan dengan baik.

Sebagai langkah terakhir, Perlu dilakukan evaluasi terhadap cabang-cabang yang ada. Evaluasi tersebut dilakukan dengan melihat data penjualan dalam 2 tahun terakhir. Data penjualan tersebut akan

* Korepondensi.

Alamat E-mail : jurnal.larik@bsi.ac.id.

Diterima 20 Mei 2022; Direvisi 20 Juli 2022; Diterima 30 Juli 2022

© 2022 Jurnal Larik.

diolah dengan metode tertentu sehingga diketahui mana cabang yang mencapai target mana yang tidak.

Oleh sebab itu sebagai bahan evaluasi, maka diperlukan algoritma yang mampu menentukan cabang tersebut apakah penjualannya baik atau tidak. Berdasarkan beberapa penelitian sebelumnya, maka dilakukan perhitungan menggunakan algoritma KMean[5] dan AHP[9][10].

Beberapa penelitian tentang clustering diantaranya dilakukan oleh Nuryaman [2] dimana algoritma kmean lebih baik dalam menentukan keanggotaan dengan nilai rasio 0,000249. Selain itu di setiap literasi adanya peregeseran titik tengah dengan nilai centroid sehingga hasil literasinya juga lebih cepat.

Clustering dilakukan dengan membagi jaringan ke dalam beberapa bagian. Pembagian ini ditujukan agar pengiriman data yang semula langsung dari SN ke Sink bisa dilakukan ke Cluster Head (CH). Pengiriman data selanjutnya dilakukan oleh CH ke Sink. CH dalam setiap cluster dilakukan penggantian secara berkala agar konsumsi energi untuk pengiriman data bisa merata ke setiap SN. Pemilihan CH dilakukan berdasarkan jumlah energi yang dimiliki. Jumlah Cluster Member (CM) setiap cluster diusahakan dalam jumlah yang berimbang sehingga tingkat pengiriman data yang dilakukan oleh CH bisa sama.

Selain itu penelitian yang dilakukan oleh Normah [3] dimana hasil penelitiannya yaitu Penerapan metode K-Means pada toko Helai, yaitu dengan cara mengelompokkan data stok baju. 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 11 artikel, yang laris ada 55 artikel dan 34 artikel untuk yang kurang laris.

Selain itu penelitian yang dilakukan oleh Yulianti [4] dengan menggunakan dua buah parameter yaitu transaksi dan jumlah penjualan dan melewati tiga kali melakukan iterasi dengan hasil iterasi satu mendapatkan nilai rasio sebesar 0,374324132, iterasi dua mendapatkan rasio 0,543018325, dan iterasi tiga mendapatkan nilai rasio yang sama dengan iterasi kedua.

K-Means merupakan salah satu metode data clustering non hierarki yang berusaha mempartisi

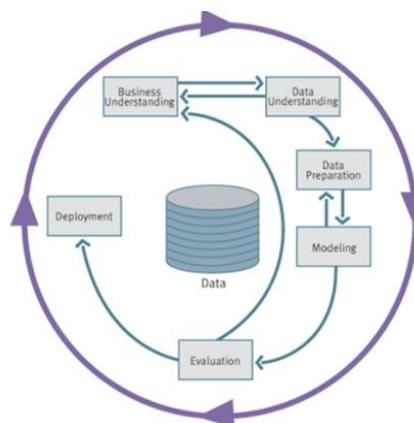
data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lainnya [7]

Selain itu terdapat penggunaan algoritma AHP[9][10] dimana hasil penelitian tersebut adalah menghasilkan model kebijakan mengenai penyakit malaria dan disusun dengan menggunakan 3 kriteria prioritas kebijakan[8].

Oleh sebab ini kami melakukan penelitian terhadap data-data penjualan yang ada dengan menggunakan algoritma k-mean dan AHP dan membandingkan hasilnya untuk melihat klasifikasi tiap toko sebagai bahan evaluasi

II. METODE PENELITIAN

Penelitian ini menggunakan model CRISP-DM (Cross-Industry Standard Proses for Data Mining) [6], yang terdiri dari 6 tahap yaitu:



Gambar 1. CRISP-DM

1. Fase Pemahaman Bisnis (*Business/Research Understanding Phase*)

Swalayan x merupakan perusahaan yang bergerak dalam usaha penjualan bahan kebutuhan sehari-hari. Swalayan x memiliki ratusan cabang di seluruh wilayah Indonesia, namun dikarenakan keterbatasan izin pemberian data, maka yang akan diteliti hanya sebanyak 10 cabang di wilayah Jakarta. 10 cabang swalayan x tersebut akan dianalisa untuk menentukan tiap-tiap toko tersebut masuk klasifikasi yang mana.

2. Fase Pemahaman Data (*Data Understanding Phase*)

Data yang diambil yaitu penjualan bulanan dari tiap swalayan selama 2 tahun terakhir yaitu tahun 2020 dan tahun 2021. Adapun swalayan yang dijadikan bahan penelitian yaitu 10 swalayan di wilayah jakarta. Adapun rinciannya adalah sebagai berikut.

TABEL I
DATA PENJUALAN SWALAYAN X

No	Cabang	2020	2021
1	Sw 1	3,389,120,100	3,104,341,300
2	Sw 2	4,899,987,600	3,408,982,400
3	Sw 3	4,346,239,100	3,164,570,400
4	Sw 4	1,898,775,500	2,322,518,600
5	Sw 5	2,452,524,000	2,440,727,600
6	Sw 6	1,898,775,500	4,067,020,400
7	Sw 7	3,792,490,600	1,563,059,100
8	Sw 8	3,792,490,600	2,979,664,200
9	Sw 9	2,452,524,000	2,667,513,400
10	Sw10	4,346,239,100	2,811,806,000

3. Fase Pengolahan Data (*Data Preparation Phase*)
Untuk mempermudah dalam perhitungan, maka dilakukan pembulatan ke angka puluhan seperti yang tampak di tabel dibawah ini

TABEL II
DATA SIAP OLAH

Data	Fitur X	Fitur Y
1	34	31
2	49	34
3	43	32
4	19.0	23.0
5	25.0	24.0
6	19.0	41.0
7	38.0	16.0
8	38.0	30.0
9	25.0	27.0
10	43.0	28.0

4. Fase Pemodelan (*Modeling Phase*)

Pada tahap ini dilakukan pemodelan menggunakan algoritma K-MEAN [5] dan AHP

5. Fase Evaluasi (*Evaluation Phase*)

Pada tahap ini pengujian model dilakukan dengan menggunakan MSE , yang nantinya hasil dari pengujian tersebut menjadi dasar pembandingan algoritma mana yang dapat digunakan.

6. Fase Penyebaran (*Deployment Phase*)

Setelah diketahui hasil klasifikasinya, maka hal tersebut disampaikan kepada swalayan xi untuk memberikan rekomendasi dalam penilaian cabang-cabang tersebut.

III. HASIL DAN PEMBAHASAN

Adapun langkah-langkah perhitungan yang dilakukan adalah sebagai berikut.

A. Perhitungan dengan Algoritma K-Mean

Pada fase ini dilakukan pemilihan titik tengah awal yang dilakukan secara acak yaitu seperti table dibawah ini

TABEL III
CENTROID AWAL

data	X	Y
C1 (4)	19.0	23.0
C2 (8)	43.0	24.0

Setelah centroid awal ditentukan, maka dilakukan perhitungan dengan rumus yang ada sehingga menghasilkan clustering seperti table dibawah ini

TABEL IV
HASIL LITERASI 1

Data	Jarak ke Centroid		Keanggotaan
	C1	C2	
1	17	11	C2
2	31	11	C2
3	25	8	C2
4	0	24	C1
5	6	18	C1
6	18	29	C1
7	20	9	C2
8	20	7	C2
9	7	18	C1
10	24	4	C2

Pada table diatas dapat dilihat ada 6 cabang yang masuk katerogri C2 dan 4 cabang masuk kategori C1. Rasio pada literasi 1 yaitu sebesar 0.0264255. Dan kemudian dilakukan perhitungan untuk mencari titik tengah yang baru yang akan digunakan pada literasi selanjutnya. Adapun setelah dilakukan perhitungan berikut adalah titik tengah yang baru

TABEL V
CENTROID AWAL

data	X	Y
C1	22.0	28.8
C2	40.8	28.5

Setelah dilakukan perhitungan dengan langkah-langkah sebelumnya, pada literasi ke 4 nilai rasio yang dihasilkan yaitu 0,03854 tidak lebih besar daripada rasio sebelumnya 0,04926. Atas dasar tersebut maka perhitungan dihentikan dan dapat dipastikan klasifikasi akhir dari tiap toko adalah

sebagai berikut. Adapun hasil perhitungan MSE adalah 0,02534

TABEL IV
HASIL LITERASI 1

Data	Anggota
1	C2
2	C2
3	C2
4	C1
5	C2
6	C1
7	C2
8	C2
9	C1
10	C2

B. Perhitungan dengan Algoritma AHP

Setelah data dimasukan kedalam table dan diolah menggunakan rumus yang ada maka didapatkan kedekatan terdekat yaitu Cabang 5 dan 9 seperti pada gambar dibawah ini

Dman	1	2	3	4	5	6	7	8	9	10
1	0	18	10	23	17	25	19	5	14	12
2	18	0	8	41	35	37	29	15	32	12
3	10	8	0	33	27	33	21	7	24	4
4	23	41	33	0	6	18	26	26	9	29
5	17	35	27	6	0	22	22	20	3	23
6	25	37	33	18	22	0	44	30	19	37
7	19	29	21	26	22	44	0	14	25	17
8	5	15	7	26	20	30	14	0	17	7
9	14	32	24	9	3	19	25	17	0	20
10	12	12	4	29	23	37	17	7	20	0

Gambar 2. Hasil pengolahan literasi 1

Setelah diketahui kedekatan awal, maka dilakukan perhitungan untuk mendapatkan kedekatan terdekat selanjutnya dengan rumus yang ada seperti yang tampak pada gambar dibawah ini

Dman	(59)	1	2	3	4	6	7	8	10
(59)	0	17	35	27	9	22	25	20	23
1	17	0	18	10	23	25	19	5	12
2	35	18	0	8	41	37	29	15	12
3	27	10	8	0	33	33	21	7	4
4	9	23	41	33	0	18	26	26	29
6	22	25	37	33	18	0	44	30	9
7	25	19	29	21	26	44	0	14	17
8	20	5	15	7	26	30	14	0	7
10	23	12	12	4	29	9	17	7	0

Gambar 3. Hasil Pengolahan literasi 2

Seperti yang tampak pada gambar 2, kedekatan selanjutnya adalah cabang 10 dan cabang 3. Selanjutnya ulangi seperti langkah sebelumnya

sehingga menghasilkan literasi ketika seperti gambar dibawah ini

Dman	(78)	(14)	2	3	5	6	9	10
(78)	0	26	29	21	21	44	24	17
(14)	26	0	41	33	16	25	13	29
2	29	41	0	8	34	37	31	12
3	21	33	8	0	26	33	23	4
5	21	16	34	26	0	23	3	22
6	44	25	37	33	23	0	20	37
9	24	13	31	23	3	20	0	19
10	17	29	12	4	22	37	19	0

Gambar 4. Hasil Pengolahan literasi 3

Pada literasi ke 3 kedekatan terjadi pada cabang 7,8 dengan 9. Selanjutnya dengan langkah seperti yang sebelumnya dilakukan maka didapatlah literasi ke 4 seperti gambar dibawah ini

Dman	(789)	(14)	2	3	5	6	10
(789)	0	26	31	23	21	44	19
(14)	26	0	41	33	16	25	29
2	31	41	0	8	34	37	12
3	23	33	8	0	26	33	4
5	21	16	34	26	0	23	22
6	44	25	37	33	23	0	37
10	19	29	12	4	22	37	0

Gambar 5. Hasil Pengolahan literasi 4

Pada literasi ke 4 kedekatan terjadi pada cabang 2 dengan 3. Selanjutnya dengan langkah seperti yang sebelumnya dilakukan maka didapatlah literasi ke 5 seperti gambar dibawah ini

Dman	(23)	(789)	(14)	5	6	10
(23)	0	31	41	34	37	12
(789)	31	0	26	21	44	19
(14)	41	6	0	16	25	29
5	34	21	16	0	23	22
6	37	44	25	23	0	37
10	12	19	29	22	37	0

Gambar 6. Hasil Pengolahan literasi 5

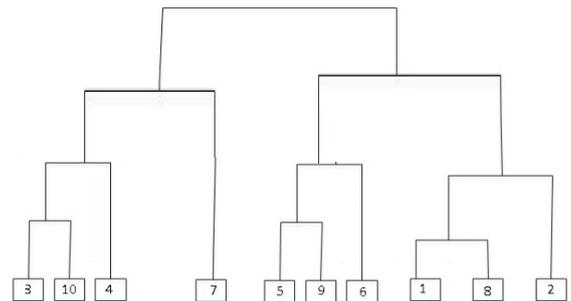
Pada literasi ke 5 kedekatan terjadi pada cabang 2,3 dengan 6. Perhitungan ini dilakukan hingga nantinya didapatkan 2 kelompok pada literasi ke 9 seperti yang tampak pada gambar dibawah ini

Dman	(218659)	(74310)
(218659)	0	44
(74310)	44	0

Gambar 7. Hasil literasi 9

Seperti yang tampak pada gambar diatas cabang yang termasuk kelompok 1 adalah cabang 7,4,3 dan 10. Untuk cabang yang termasuk kelompok 2 adalah cabang 6,5,9 2,1 dan 8. Pengukuran performance dilakukan menggunakan Mean Square Error, dan menghasilkan nilai MSE sebesar 0,03342.

Adapun untuk melihat pembentukan kelompok tersebut dapat digambarkan pada gambar di bawah ini



Gambar 8. Gambaran kedekatan setiap cabang swalayan

IV. KESIMPULAN DAN SARAN

Berdasarkan hasil perhitungan menggunakan Algoritma Kmean yang masuk C1 berjumlah 4 yaitu cabang 4,5,6, dan 9. Dan yang masuk C2 berjumlah 6 yaitu cabang 1,2,3,7,8, dan 10. Sedangkan perhitungan menggunakan Algoritma AHP yang masuk kelompok 1 berjumlah 4 yaitu 3,4,7 dan 10. Dan yang masuk kelompok 2 yaitu 1,2,5,6,8 dan 9.

Rasio MSE Algoritma Kmean adalah 0,02534, sedangkan Algoritma AHP adalah 0,03342. Sehingga dapat disimpulkan Algoritma Kmean lebih baik karena memiliki rasio yang lebih baik daripada Algoritma AHP. Dan Cabang 4,5,6 dan 9 direkomendasikan untuk diberikan evaluasi karena masuk kategori C1.

V. REFERENSI

[1] Butsianto and N. T. Mayangwulan, "Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering," J. Nas. Komputasi dan

- Teknol. Inf., vol. 3, no. 3, pp. 187–201, 2020, doi: 10.32672/jnkti.v3i3.2428
- [2] Nuryaman, Y.(2017). "Komparasi Algoritma Kmean dan Ahc untuk Klasifikasi Curah Hujan di Indonesia." *Ikraith Informatika*, vol. 2, no. 2, 1 Jul. 2018, pp. 70-75.
- [3] Normah, N., Nurajizah, S., & Salbinda, A. (2021). Penerapan Data Mining Metode K-Means Clustering untuk Analisa Penjualan Pada Toko Fashion Hijab Banten. *JITK AMIK BSI*, 22(2)
- [4] Yulianti, Y., Utami, D. Y., Hikmah, N., & Hasan, F. N. (2019). Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab. *Jurnal Pilar Nusa Mandiri*, 15(2), 241–246. <https://doi.org/10.33480/pilar.v15i2.650>
- [5] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [6] I. Parlina, A. P. Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Assessment Center. *Journal of Computer Engineering, System and Science*, 2018; 3(1): 87-93.
- [7] 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>
- [8] Armaita, Dedi , H., Eri , B., Indang , D., & Iswandi, U. (2020). Policy Model of Community Adaptation using AHP in the Malaria Endemic Region of Lahat Regency - Indonesia. *International Journal of Management and Humanities (IJMH)*, 44-48.
- [9] Deny , S., Agus , R., & Yushar , K. (2020). ANALISIS PENENTUAN TIPE FONDASI PILAR JEMBATAN DENGAN MENGGUNAKAN METODE ANALYTICAL HIERARCHY PROCESS (AHP) (Studi Kasus: Pembangunan Jembatan Walahar Kecamatan Ciampel, Kabupaten Karawang)
- [10] Armaita, Dedi , H., Eri , B., Indang , D., & Iswandi, U. (2020). Policy Model of Community Adaptation using AHP in the Malaria Endemic Region of Lahat Regency - Indonesia. *International Journal of Management and Humanities (IJMH)*, 44-48.