

PENERAPAN ALGORITMA K-MEANS UNTUK MENGELOMPOKAN MENU TERLARIS DI RUMAH MAKAN FAHJRI CHANIAGO

Fahjri Septi Uthama Chania¹, Yuli Komalasari²

^{1,2}Teknologi Informatika
Universitas Bina Sarana Informatika
Jl. Kramat Raya no.98, Jakarta Pusat 10450
E-mail: 117200631@bsi.ac.id, yuli.yks@bsi.ac.id

Abstrak

Perkembangan bisnis Usaha Mikro Kecil dan Menengah (UMKM) dalam bidang makanan merupakan salah satu usaha yang dibutuhkan masyarakat, karena hanya dengan makan manusia dapat melangsungkan hidupnya. Dengan alasan itu banyak pengusaha yang mulai bermunculan dalam bisnis kuliner yang pada akhirnya berdampak semakin ketatnya persaingan. Salah satu industri makanan yang memiliki daya saing tinggi adalah rumah makan padang. Menu di rumah makan ini bisa di cluster untuk mengetahui apa saja yang menjadi menu terlaris di rumah makan tersebut. Salah satu algoritma clustering adalah k-means. Beberapa penelitian serupa dilakukan menggunakan k-means dengan menghitung jarak menggunakan rumus *euclidean distance*. Penelitian ini akan men cluster daftar menu rumah makan Fahjri Chaniago menjadi 3 cluster yaitu cluster kriteria menu terlaris, kriteria menu menengah dan kriteria menu terendah. Dari hasil perhitungan clustering tersebut menghasilkan menu terlaris pada cluster 0 yang memiliki 3 anggota menu yaitu Tunjang, kikir, Gulai kakap. Untuk kriteria menu menengah memiliki 23 anggota menu, serta kriteria menu terendah memiliki 14 anggota menu

Kata kunci: *K-Means Clustering, Data Mining, Masakan Padang*

Abstract

The development of Micro, Small and Medium Enterprises (MSMEs) in the food sector is one of the businesses that society needs, because only by eating can humans survive. For this reason, many entrepreneurs are starting to emerge in the culinary business, which in turn has the impact of increasingly fierce competition. One of the food industries that has high competitiveness is Padang restaurants. The menu at this restaurant can be clustered to find out what is the best-selling menu at that restaurant. One of the clustering algorithms is k-means. Several similar studies were carried out using k-means by calculating distances using the Euclidean distance formula. This research will cluster the Fahjri Chaniago restaurant menu list into 3 clusters, namely the best-selling menu criteria cluster, medium menu criteria and lowest menu criteria. From the results of the clustering calculations, it produces the best-selling menu in cluster 0 which has 3 menu members, namely Tunjang, kikir, snapper curry. For menu criteria, the middle menu has 23 menu members, and the lowest menu criteria has 14 menu members

Keywords: *K-Means Clustering, Data Mining, Padang Food*

1. PENDAHULUAN

Perkembangan bisnis Usaha Mikro Kecil dan Menengah (UMKM) dalam bidang makanan merupakan salah satu usaha yang dibutuhkan masyarakat, karena hanya dengan makan manusia dapat melangsungkan hidupnya. Dengan alasan itu banyak pengusaha yang mulai bermunculan dalam bisnis kuliner yang pada akhirnya berdampak semakin ketatnya persaingan. Penerapan data mining pada penelitian ini adalah untuk clustering menu terlaris di rumah makan Fahjri Chaniago. Penelitian ini menggunakan jenis data kuantitatif, yaitu berupa angka atau nominal data yang dapat dihitung. Sumber data yang digunakan merupakan data primer karena data yang didapat berasal dari sumber asli atau pertama. Data yang dibutuhkan dalam penelitian ini adalah data penjualan sebelumnya, pengumpulan data berupa data set yang didapat kemudian dianalisis dan diterapkan untuk mendapatkan hasil kesimpulan penjualan.

Tujuan penelitian ini yaitu Menganalisis apakah terdapat pengaruh dengan diterapkannya metode clustering pada Rumah Makan Fahjri Chaniago. Clustering adalah proses pengelompokan benda serupa ke dalam kelompok yang berbeda, sehingga data dalam setiap subset memiliki arti yang bermanfaat. Algoritma clustering terdiri dari dua bagian yaitu secara hirarkis dan secara partisional. Algoritma hirarkis menemukan cluster secara berurutan dimana cluster ditetapkan sebelumnya, sedangkan algoritma partisional menentukan semua kelompok pada waktu tertentu. Clustering juga bisa dikatakan suatu proses dimana mengelompokkan dan membagi pola data menjadi beberapa jumlah set sehingga akan membentuk pola yang serupa dan dikelompokkan pada cluster yang sama serta memisahkan diri dengan pola yang berbeda. (Indraputra & Fitriana, t.t., 2020)

Menurut (Nugroho dkk., 2022) langkah-langkah melakukan clustering dengan metode K-means adalah sebagai berikut:

- 1) Tentukan jumlah cluster K.
- 2) Inisialisasi K pusat cluster. Ini dapat dilakukan dengan berbagai cara, namun cara yang paling banyak digunakan adalah dengan menentukan secara random. Pusat cluster diberi nilai awal dengan angka-angka random
- 3) Alokasikan semua data atau objek ke cluster terdekat. kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut demikian juga kedekatan suatu data ke cluster ditentukan jarak antara data dengan pusat cluster. Untuk menghitung jarak semua data masuk ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

2. METODE PENELITIAN



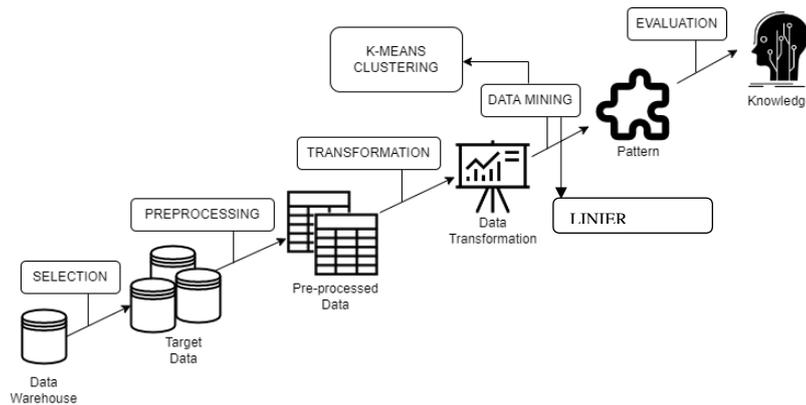
Sumber: (Prasetiani & Rochmawati, 2022)

Gambar 1
Tahapan Penelitian

Gambar 1 menunjukkan rangkaian tahapan penelitian yang dilakukan. Instrumen untuk mengumpulkan data yang digunakan dalam penelitian ini adalah angket, penyebaran angket kepada pihak-pihak yang mempunyai peranan dengan masalah yang diteliti. Penyebaran kuesioner secara tidak langsung yaitu dibantu dengan layanan fitur google form sebanyak 100 responden. Populasi dalam penelitian ini adalah seluruh pelanggan rumah makan Fahjri Chaniago sebanyak 1000 pembeli dalam 1 bulan. Teknik pengambilan sampel yang digunakan dalam penelitian ini adalah sampling random atau sampling probabilitas yaitu cara pengambilan sampel dengan semua obyek atau elemen populasi yang memiliki kesempatan yang sama untuk dipilih sebagai sampel. Dalam menentukan ukuran sampel, peneliti menggunakan rumus slovin sebagai berikut :

$$n = \frac{N}{1 + N(e)^2}$$

2.1. Pengolahan Data Mining



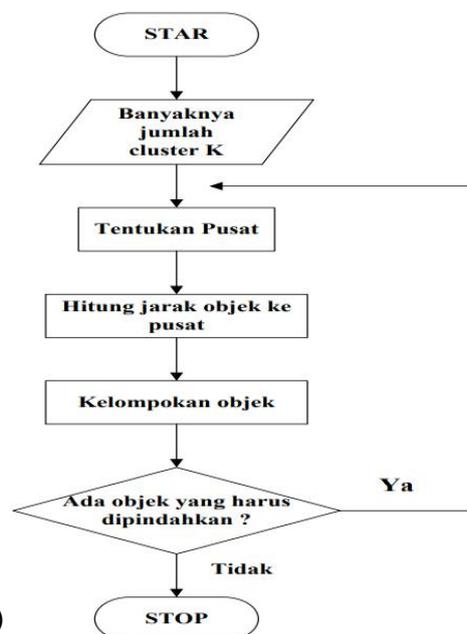
Sumber: (Asyifa Tasya, 2020)

Gambar 2
 Tahapan proses *Knowledge Discovery In Database* (KDD)

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Data mining (gambar 2) digunakan untuk mengekstrasi pengetahuan dari kumpulan data menggunakan teknik statistik dan matematika. Teknik statistik dan matematika digunakan untuk mengetahui keakuratan setiap proses data mining sehingga data ekstrasi dapat diolah menjadi informasi yang berguna.(Komalasari dkk., 2024)

2.2. Analisa Clustering Dengan Algoritma K-Means

Gambar dibawah merupakan gambar flowchart dari algoritma K-means yang digunakan dan menggambarkan langkah-langkah dalam algoritma k-means dengan anggapan bahwa input adalah jumlah data set sebanyak n data(Samudra et al., 2021) jumlah inialisasi centroid k=3 sesuai dengan penelitian



Sumber : (Asyifa Tasya, 2020)

Gambar 2
 FlowSchart Proses K-Means

2.3. Uji Validitas

Menurut (Erika et al., 2021) Menyatakan bahwa suatu pertanyaan dikatakan valid apabila tingkat signifikasinya berada kurang dari angka 0,05. Uji validasi dilakukan dengan cara membandingkan nilai r hitung dengan r tabel untuk tingkat signifikan 5 persen. Persyaratan dikatakan valid jumlah sampel jika :

- a) $r \text{ hitung} > r \text{ tabel}$ maka pertanyaan atau indikator tersebut dinyatakan valid
- b) $r \text{ hitung} < r \text{ tabel}$ maka pertanyaan atau indikator tersebut dinyatakan tidak valid

2.4. Uji Reabilitas

Menurut (Erika et al., 2021) Uji Reabilitas merupakan alat untuk suatu kuesioner yang merupakan indikator dari variabel atau konstruk Suatu konstruk atau variabel dikatakan reliable jika memberikan nilai *Cornbach Alpha* $> 0,60$.

2.5. Uji Normalitas

Menurut (Erika et al., 2021) Uji normalitas bertujuan untuk mengetahui apakah data penelitian yang didapatkan berdistribusi normal atau mendekati normal. Apabila nilai probabilitasnya > 0.05 maka data tersebut dinyatakan berdistribusi normal, begitu pula sebaliknya.

3. HASIL DAN PEMBAHASAN

Responden dalam penelitian adalah pelanggan rumah makan Fahjri Chaniago berjumlah 100 orang. Dan menyebarkan kuesioner dengan 10 soal Analisis Kuantitatif terdiri dari pengujian instrument Uji validitas, Uji reabilitas, dan Uji Normalitas dan pengolahan data menggunakan metode K-means dilakukan dengan menggunakan bantuan perangkat lunak (*software*) SPSS (*Statistical Product and Seivicei Aoluition*) Versi 25.0.

3.1. Tahap Seleksi Data

Data yang akan diolah oleh penulis adalah berupa data daftar menu dengan jumlah transaksi dan harga menu selama periode waktu tertentu adalah 40 titik data yang akan dimasukkan ke dalam kategori menu penelitian ini. Periode transaksi selama 1 bulan, atau bulan Mei 2024, adapun data yang di olah sebagai berikut.

Tabel 1
Kategori menu penelitian

NO	Nama Menu	Harga	Transaksi
1.	Ayam Bakar	17000	776
2.	Ayam Goreng	17000	689
3.	Ayam opor	17000	695
4.	Ayam balado	17000	688
5.	Rendang	17000	730
6.	Tunjang	28000	787
7.	Kikil	23000	550
8.	Ati ampla	17000	543
9.	Ati sapi	18000	532
10.	Babat	17000	572
11.	paru	18000	345
12.	limpa	18000	386

13.	Gulai Cincang	17000	494
14.	Telor dadar	16000	640
15.	Telor bulat	15000	633
16.	Lele goreng	16000	670
17.	Mujair goreng	16000	578
18.	Udang goreng	17000	367
19.	Tongkol cabai merah	16000	482
20.	Tongkol cabai hijau	16000	422
21.	Gembung goreng	16000	401
22.	Gembung bakar	16000	620
23.	Gembung asam padeh	16000	534
24.	Bawal bakar	16000	462
25.	Ikan mas	17000	340
26.	Gulai Kakap	26000	755
27.	Tahu goreng	1000	400
28.	Tempe goreng	1000	400
29.	Perkedel	2000	480
30.	Terong	2000	380
31.	Jengkol balado	4000	685
32.	Sayur nangka	5000	550
33.	Kerupuk peyek	5000	293
34.	Kerupuk ubi	2000	328
35.	Kerupuk putih	2000	321
36.	Es teh manis	3000	305
37.	Es jeruk	5000	313
38.	Teh pucuk	5000	278
39.	Teh botol	5000	287
40.	Air mineral	5000	302

Sumber : Hasil olah data, 2024

3.2. Menentukan Jumlah Cluster Dan Nilai Centroid Awal

Berdasarkan pada tabel 1 dilakukan perhitungan dengan menggunakan algoritma K-means yang diawali langkah pertama yaitu penetapan jumlah titik pusat cluster atau centroid (K) yaitu 3, kemudian langkah kedua yaitu pemilihan centroid diambil secara acak pada data yang terdapat pada tabel 2

Tabel 2
 Centroid Awal

Centroid	Menu	harga	Transaksi
C1	Telor Bulat	15000	633
C2	Ikan Mas	17000	340
C3	Rendang	17000	730

Sumber : Hasil olah data, 2024

3.3. Menghitung Jarak Setiap Jumlah

Pada langkah ini setelah menentukan titik pusat cluster yang perlu dilakukan adalah menghitung jarak antara data dengan pusat cluster dengan cara menggunakan persamaan *Euclidean*. Berikut merupakan perhitungan jarak pada setiap data pada C1

3.4. Melakukan Iterasi Ke-1

Jarak ini menunjukkan bahwa data berada dalam satu kelompok dengan pusat dengan pusat cluster terdekat dengan membandingkan hasil cluster dan diambil yang terkecil. Jika nilai terendah terdapat di cluster 1 (C1) maka masuk ke dalam cluster 1 dan seterusnya. Hasil dari pengelompokkan iterasi 1 dapat dilihat pada tabel berikut.

Tabel 3
 Hasil cluster Iterasi 1

Cluster	Nilai
C1	20
C2	10
C3	10

Sumber : Hasil olah data, 2024

3.5. Menentukan Centroid Baru

Sistem akan menjalankan kembali perhitungan menggunakan nilai titik pusat (centroid) terbaru. Sampai kondisi untuk mengakhiri proses k-means clustering terpenuhi, yaitu ketika hasil pengelompokkan dan nilai centroid yang dibuat tetap tidak berubah dari hasil perhitungan sebelumnya, terus ulangi langkah kedua hingga pada iterasi data sudah tidak mengalami perubahan lagi atau hasil pengelompokkan sama dengan centroid sebelumnya. Pada pengujian ini, iterasi ke-2 hingga iterasi ke-5 mengalami perubahan pada centroid dan hasil pengelompokkan. Pada pengujian iterasi ke-6 data tidak mengalami perubahan pada centroid dan hasil pengelompokkan. Berikut tabel dibawah merupakan hasil akhir clustering

Tabel 4
 Pengelompokkan Iterasi Akhir

NO	C1	C2	C3	JARAK TERDEKAT	KELAS
1	13648.60	416.01	8667.02	416.01	2
2	13646.35	375.40	8666.67	375.40	2
3	13646.49	377.70	8666.67	377.70	2
4	13646.33	375.03	8666.67	375.03	2
5	13647.34	392.67	8666.73	392.67	2
6	24646.22	11350.35	2335.06	2335.06	3
7	19643.59	6347.83	2670.73	2670.73	3
8	13643.83	347.86	8668.04	347.86	2
9	14643.64	1347.92	7668.45	1347.92	2
10	13644.21	348.67	8667.57	348.67	2
11	14642.90	1363.00	7674.76	1363.00	2
12	14642.86	1357.50	7672.99	1357.50	2
13	13643.33	351.96	8669.05	351.96	2
14	12645.53	658.66	9666.84	658.66	2

15	11645.60	1654.37	10666.86	1654.37	2
16	12646.18	663.53	9666.71	663.53	2
17	12644.41	652.87	9667.40	652.87	2
18	13642.86	392.00	8672.96	392.00	2
19	12643.27	655.48	9669.06	655.48	2
20	12642.93	664.19	9670.59	664.19	2
21	12642.87	668.49	9671.21	668.49	2
22	12645.13	656.16	9666.98	656.16	2
23	12643.79	652.32	9668.05	652.32	2
24	12643.12	657.79	9669.53	657.79	2
25	13642.92	405.16	8674.03	405.16	2
26	22645.96	9350.12	338.28	338.28	3
27	2357.23	15652.87	24668.46	2357.23	1
28	2357.23	15652.87	24668.46	2357.23	1
29	1360.81	14652.33	23667.66	1360.81	1
30	1357.14	14653.13	23668.79	1357.14	1
31	711.48	12652.92	21666.67	711.48	1
32	1651.61	11652.17	20667.19	1651.61	1
33	1645.17	11654.96	20670.62	1645.17	1
34	1358.14	14653.82	23669.55	1358.14	1
35	1358.43	14653.93	23669.66	1358.43	1
36	364.96	13654.33	22670.06	364.96	1
37	1644.23	11654.54	20670.24	1644.23	1
38	1646.03	11655.30	20670.92	1646.03	1
39	1645.50	11655.09	20670.74	1645.50	1
40	1644.71	11654.77	20670.45	1644.71	1

Sumber : Hasil olah data, 2024

3.6. Hasil Kelompok (Clustering)

Setelah didapatkan hasil akhir dari proses iterasi K-means clustering diperoleh hasil pengelompokan pada masing masing cluster. Berikut tabel hasil cluster Menu Rumah Makan Fahjri Chaniago

Tabel 5
 Hasil Clustering

Cluster	Jumlah	Anggota Menu
Cluster 0	3	Tunjang
		Kikil
		Gulai Kakap
Cluster 1	23	Ayam Bakar
		Ayam Goreng
		Ayam opor
		Ayam balado

		Rendang
		Ati ampla
		Ati sapi
		Babat
		paru
		limpa
		Gulai Cincang
		Telor dadar
		Telor bulat
		Lele goreng
		Mujair goreng
		Udang goreng
		Tongkol cabai merah
		Tongkol cabai hijau
		Gembung goreng
		Gembung bakar
		Gembung asam padeh
		Bawal bakar
		Ikan mas
Cluster 2	14	Tahu goreng
		Tempe goreng
		Perkedel
		Terong
		Jengkol balado
		Sayur nangka
		Kerupuk peyek
		Kerupuk ubi
		Kerupuk putih
		Es teh manis
		Es jeruk
		Teh pucuk
		Teh botol
		Air mineral

Sumber : Hasil olah data, 2024

- Indraputra, R. A., & Fitriana, R. (2020). K-Means clustering data COVID-19. *Jurnal Teknik Industri*, 10(3), 275–282.
- Komalasari, Y. & Rahmah, P. (2024). Penerapan Algoritma C4.5 Untuk Menentukan Kepuasan Pengguna Aplikasi E-Open Study Kasus : Kelurahan Jati Makmur
- Marvell Geraldine, Y., & Susanti, A. (2021). *POINT PENGARUH MEREK, KUALITAS PRODUK, MEDIA SOSIAL DAN HARGA TERHADAP MINAT BELI KONSUMEN PADA PRODUK BRAND WARDAH*. <https://ejournals.umma.ac.id/indeks.php/point>
- Megawati. (2022). *STRATEGI MEMPERTAHANKAN EKSITENSI DAN PENDAPATAN RUMAH MAKAN PONDOK BAMBU MOROTUMAN PASCA COVID-19 DESA HASANAH KEC MAPPEDECENG KAB LUWU UTARA*.
- Nugroho, M. R., Hendrawan, I. E., & Purwantoro, P. P. (2022). Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI. *Nuansa Informatika*, 16(1), 125–133.
- Nurlina. (2020). *ANALISIS BAURAN PEMASARAN PADA MASA NEW NORMAL DI PRAKTIK KEPERAWATAN MANDIRI KOTA TASIKMALAYA MARKETING MIX ANALISYS IN NEW NORMAL ERA IN INDEPENDENT NURSING PRACTICES TASIKAMLAYA CITY 1*.
- Prasetiani, S. D., & Rochmawati, N. (2022). Penerapan Data Mining Untuk Clustering Menu Favorit Menggunakan Algoritma K-Means (Studi Kasus Kedai Expo). *Journal of Informatics and Computer Science*, 03.
- Samudra, F. P., Pudjiantoro, T. H., & Santikarama, I. (2021). Klasterisasi Tingkat Penjualan Produk Menggunakan Metode K-Means Untuk Penerapan Konsep Down-Selling: Studi Kasus pada Artch Indonesia. *SNIA (Seminar Nasional Informatika Dan Aplikasinya)*, 5, B9-15.