

## IMPLEMENTASI ALGORITMA *PARTITIONING AROUND MEDOIDS* DALAM PENGELOMPOKAN RESTORAN

Tito Bimantoro<sup>1</sup>, Anindya Khrisna Wardhani<sup>2</sup>

<sup>1</sup>Universitas Dian Nuswantoro, <sup>2</sup>Universitas Diponegoro  
Email: <sup>1</sup>bimantoro25@gmail.com, <sup>2</sup>nindywardhani77@gmail.com

(Naskah masuk: 10 Desember 2020, diterima untuk diterbitkan: 29 Desember 2020)

### Abstrak

Banyaknya data yang hanya sebatas memberikan grafik atau statistik mengenai data restoran dapat dimanfaatkan untuk penggalian data menggunakan teknik *data mining*. Informasi yang dapat digali dari data restoran adalah pengelompokan kategori restoran berdasarkan jumlah menu yang terdapat pada restoran tersebut. Penelitian ini menerapkan teknik *data mining* menggunakan metode *clustering partitioning around medoids (k-medoids)* untuk menghasilkan informasi mengenai pengelompokan kategori restoran. Algoritma *clustering Partitioning Around Medoids (PAM)* atau disebut juga *K-Medoids* adalah algoritma untuk menemukan *k cluster* dalam *n* objek dengan pertama kali secara arbitrarly menemukan wakil dari objek (*medoid*) untuk tiap-tiap *cluster*. Pengelompokan dalam penelitian ini adalah *appetizer, soup, main course, dessert, minuman dan snack*. Kemudian hasil tersebut dapat dijadikan bahan atau dasar dari konsumen untuk memilih restoran berdasarkan kategori apa yang ingin konsumen makan. Berdasarkan data yang diperoleh, dihasilkan *cluster* yang merepresentasikan *appetizer (cluster\_2)* 17 item, *soup (cluster\_1)* 7 item, *main course (cluster\_4)* 51 item, *dessert (cluster\_3)* 11 item, minuman (*cluster\_5*) 6 item dan *snack (cluster\_0)* 13 item.

**Kata kunci:** *Data Mining, K- Medoids, Clustering*

## IMPLEMENTATION OF PARTITIONING ALGORITHM AROUND MEDOIDS IN CLUSTERING RESTAURANT

### Abstract

The amount of data that is only limited to providing graphs or statistics regarding restaurant data can be used for data mining using data mining techniques. The information that can be extracted from restaurant data is the grouping of restaurant categories based on the number of menus in the restaurant. This study applies data mining techniques using clustering partitioning around medoids (k-medoids) to generate information about grouping restaurant categories. The clustering algorithm for Partitioning Around Medoids (PAM) or also known as K-Medoids is an algorithm for finding k clusters in n objects by arbitrarily finding the representative of the object (medoid) for each cluster. The grouping in this research is appetizer, soup, main course, dessert, drink and snack. Then these results can be used as ingredients or the basis for consumers to choose a restaurant based on what category consumers want to eat. Based on the data obtained, a cluster is produced that represents an appetizer (cluster\_2) 17 items, soup (cluster\_1) 7 items, main course (cluster\_4) 51 items, dessert (cluster\_3) 11 items, drinks (cluster\_5) 6 items and snacks (cluster\_0) 13 item.

**Keywords:** *Data Mining, K- Medoids, Clustering*

### 1. PENDAHULUAN

Keanekaragaman makanan khas pada suatu daerah membuat banyak para wisatawan lokal maupun luar daerah berkunjung untuk menikmati wisata kuliner (Benny, 2014). Banyaknya data rumah makan dapat dimanfaatkan untuk menjadi sebuah informasi yang dapat digali. Salah satu teknik yang mampu menggali informasi tersembunyi dari kumpulan data multidimensi yang telah diperoleh adalah teknik *data mining*. Salah satu teknik *data mining* yang cukup terkenal yaitu *clustering*.

*Clustering* adalah metode yang digunakan dalam *data mining* yang cara kerjanya mencari dan mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh (Ong dkk., 2013).

Metode ini dikembangkan dengan menentukan struktur pengelompokan diantara objek yang akan dikelompokkan. Analisis cluster dimulai dari metode hierarchial yang secara garis besar membentuk sebuah tree diagram yang biasa disebut dengan dendogram yang mendeskripsikan pengelompokan

berdasarkan jarak, *graph-theoretic* melihat objek sebagai *node* pada *network* terboboti, *mixture models* mengasumsikan suatu objek dihasilkan dari skala data yang berbeda-beda, *partitional* lebih dikenal dengan metode *non-hierarchy* termasuk didalamnya adalah metode *K-Medoids cluster*. Algoritma *clustering Partitioning Around Medoids(PAM)* atau disebut juga *K-Medoids* adalah algoritma untuk menemukan *k cluster* dalam *n* objek dengan pertama kali secara arbitrary menemukan wakil dari objek (*medoid*) untuk tiap-tiap *cluster* (Velmurugan, 2010)

Pada penelitian ini, algoritma *Partitioning Around Medoids* digunakan dalam pengelompokan rumah makan berdasarkan kriteria menu yang terdapat pada restoran tersebut. Pengelompokan dibagi menjadi 6 cluster yaitu *appetizer*, *soup*, *main course*, *dessert*, minuman dan *snack*.

## 2. METODE

### 2.1 Metode Pengumpulan dan Analisis Data

Dalam kegiatan pengumpulan data untuk penelitian ini digunakan metode pengumpulan studi pustaka yang mana pada metode ini kegiatan dilakukan adalah mempelajari, mencari dan mengumpulkan data yang berhubungan dengan penelitian ini. Data yang digunakan dalam pengelompokan restoran diperoleh dari observasi dan wawancara langsung dengan pemilik restoran. Sedangkan studi pustaka dengan cara mengumpulkan literatur pendukung penelitian, baik dari buku referensi maupun dari internet.

Data penelitian yang digunakan untuk penelitian adalah data restoran yang terdapat di Kota Semarang sebanyak 105 data.

### 2.2 Algoritma clustering Partitioning Around Medoids

Pada umumnya kinerja metode Algoritma *clustering Partitioning Around Medoids* secara berurutan adalah sebagai berikut :

1. Tentukan *k* (jumlah *cluster*) yang diinginkan
2. Pilih secara acak *medoid* awal sebanyak *k* dari *n* data
3. Hitung jarak masing-masing obyek ke *medoid* sementara, kemudian tandai jarak terdekat obyek ke *medoid* dan hitung totalnya.
4. Lakukan iterasi *medoid*
5. Hitung total simpangan (S)  
Jika *a* adalah jumlah jarak terdekat antara obyek ke *medoid* awal, dan *b* adalah jumlah jarak terdekat antara obyek ke *medoid* baru, maka total simpangan adalah  $S = b - a$   
Jika  $S < 0$ , maka tukar obyek dengan data untuk membentuk sekumpulan *k* baru sebagai *medoid*

6. Ulangi langkah 3 sampai 5 dan hentikan jika sudah tidak terjadi perubahan anggota *medoid*

## 3. HASIL DAN PEMBAHASAN

### 3.1. Sample Data

Data penelitian yang diperoleh dari observasi dan wawancara restoran di Kota Semarang kemudian diolah dengan cara difilter hanya untuk restoran yang memiliki menu makanan khas dan dihitung jumlah menu pada setiap kategorinya, sehingga dapat digunakan untuk *clustering*. Dari hasil pengolahan data, diperoleh data restoran yang memiliki makanan khas di Kota Semarang sebanyak 105 data yang akan dikelompokkan ke dalam enam cluster yaitu *appetizer*, *soup*, *main course*, *dessert*, minuman dan *snack*. Pengelompokan tersebut berdasarkan atribut jumlah menu yang tersedia pada restoran tersebut, yang kemudian atribut tersebut akan diolah menggunakan algoritma *Partitioning Around Medoids*.

Tabel 1. Sample Data

Nama Restoran	Jumlah Menu					
	Appetizer	Soup	Main Course	Dessert	Snack	Drink
Bandeng juwana	4	0	8	1	10	0
Elrina Pandanaran	0	0	2	0	0	0
bandeng presto	5	0	3	0	2	0
Bandeng presto bonafide	0	0	2	0	0	0
bandeng presto gemari	0	0	1	0	0	0
bandeng presto kampoeng	0	0	2	0	0	0
semarang bandeng presto suyudono	0	0	0	9	0	0
Excellent Bakery Hineni	0	0	4	0	0	0
Spesial gongso	0	0	0	0	0	9
Jamu Jun Lamper	0	0	0	0	0	9

### 3.2. Implementasi Sistem

Pada penelitian ini menggunakan *tool rapid miner*. Berikut adalah pengolahan data dengan menggunakan *k-medoids* pada *rapid miner*. Sample data yang digunakan dapat dilihat pada Gambar 1.

Row No.	Resto	Appetizer	Soup	Main Course	Dessert	Snack	Drink
1	Banding Juwana Elhina Pandanaran	4	0	8	1	10	0
2	banding presto	0	0	2	0	0	0
3	Banding presto bonafide	5	0	3	0	2	0
4	banding presto gemani	0	0	2	0	0	0
5	banding presto kampoeng semarang	0	0	1	0	0	0
6	banding presto sujudno	0	0	2	0	0	0
7	Es puffer cong lik cabang Ahmad Dahl	0	0	0	10	0	0
8	Es puffer cong lik cabang Dargo	0	0	0	16	0	0
9	Excellent Bakery	0	0	0	9	0	0
10	Hinini Spesial gongso	0	0	4	0	0	0
11	Jamu Jun Lempur	0	0	0	0	0	9
12	Jamu Jun Pandhowo	0	0	0	0	0	9
13	Jamu Jun Sampangan	0	0	0	0	0	9
14	Kabon&Kedai Aisyah	0	8	0	0	0	0
15	Kue Moad Gemini	0	0	0	0	7	0
16	Lekker pamo	0	0	0	0	10	0
17	Lorong tahu gimbal Pak Polo	0	0	1	0	0	0
18	Lumpia design dik.meme	9	0	0	0	0	0

Gambar 1. Input Data

Dalam pemrosesan pemodelan *Clustering K-Medoids* pada *Rapid Miner* dapat dilihat pada Gambar 2.



Gambar 2. Pemodelan *Clustering K-Medoids* pada *Rapid Miner*

Dengan menggunakan pemodelan *k-medoids clustering* seperti gambar 1 diatas, dengan inisialisasi jumlah *cluster* sebanyak enam buah, maka didapatkan hasil *cluster* yang terbentuk adalah enam, sesuai dengan pendefinisian nilai *k* dengan jumlah setiap *cluster* dengan total 105 data sebagai berikut :

1. *cluster\_0* = 13 items
2. *cluster\_1* = 7 items
3. *cluster\_2* = 17 items
4. *cluster\_3* = 11 items
5. *cluster\_4* = 51 items
6. *cluster\_5* = 6 items

Data hasil *clustering* pada *rapid miner* dapat dilihat pada gambar 3.

Cluster Model	
Cluster 0:	13 items
Cluster 1:	7 items
Cluster 2:	17 items
Cluster 3:	11 items
Cluster 4:	51 items
Cluster 5:	6 items
Total number of items:	105

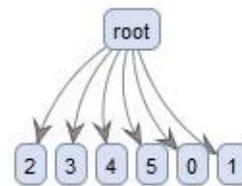
Gambar 3. Hasil *Clustering*

Hasil perhitungan antara jarak *cluster* dan *centroid* menggunakan *k-medoids* dapat dilihat pada gambar 4.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5
Resto	103	93	88	104	79	96
Appetizer	0	0	8	0	0	0
Soup	0	8	0	0	0	0
Main Course	0	0	0	0	6	0
Dessert	0	0	0	9	0	0
Snack	10	0	0	0	0	0
Drink	0	0	0	0	0	9

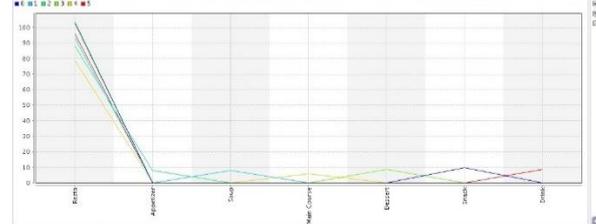
Gambar 4. Hasil perhitungan antara jarak *cluster* dan *centroid*

Hasil *root cluster* pada *clustering k-medoids* menggunakan *rapid miner* dapat dilihat pada gambar 5.



Gambar 5. Hasil *root cluster*

Grafik hasil data *cluster* dapat dilihat pada gambar 6.



Gambar 6 Grafik data hasil *Clustering K-Medoids*

## 4. KESIMPULAN DAN SARAN

### 4.1. Kesimpulan

Tumpukan data restoran dapat digunakan untuk menentukan konsistensi data kategori menggunakan teknik *data mining* yang mampu menggali informasi tersembunyi dari kumpulan data multidimensi yang telah diperoleh, selain itu pengekstrakan data yang terhubung dengan data lain juga dapat dilakukan oleh teknik *data mining* ini. Salah satu teknik *data mining* yang cukup terkenal yaitu *clustering* dan metode yang cukup populer dalam teknik *data mining* ini adalah metode *k-medoids*. *Partitioning around medoids (PAM)* atau lebih dikenal dengan algoritma *k-Medoids* merupakan salah satu algoritma *clustering data mining*. Dengan data uji berupa jumlah menu yang telah dikelompokkan kedalam kategori jenis makanan sehingga dapat menghasilkan kategori dari sebuah restoran sebanyak 105 data, dihasilkan *cluster* yang merepresentasikan *appetizer (cluster\_2)* 17 items, *soup (cluster\_1)* 7 items, *main course (cluster\_4)* 51 items, *dessert (cluster\_3)* 11 items, *minuman (cluster\_5)* 6 items dan *snack (cluster\_0)* 13 items. Algoritma *k-medoids* telah bekerja dengan baik untuk menentukan setiap objek pada dataset pada setiap *cluster*. Berdasarkan uraian masalah di atas, maka dapat dirumuskan kesimpulan inisialisasi

jumlah *cluster* sebanyak enam buah sesuai dengan pendefinisian nilai *k*.

#### 4.2. Saran

Penelitian ini dapat dikembangkan menjadi aplikasi maupun dibuat menggunakan bahasa pemrograman. Pengembangan penelitian selanjutnya dapat dikembangkan menggunakan perbandingan algoritma lainnya. Pada Penelitian ini, data yang digunakan hanya data restoran yang memiliki makanan khas, pada penelitian selanjutnya dapat dikembangkan jumlah data yang digunakan lebih secara luas.

#### DAFTAR PUSTAKA

- A. K. Wardhani, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Kajen Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- T. Velmurugan, "Efficiency of k-Means and K-Medoids Algorithms for Clustering Arbitrary DataPoints," *International Journal Computer Technology & Applications*, vol. 3, 2012.
- A. K. Wardhani, C. E. Widodo, J. E. Suseno, "Information System for Culinary Product Selection Using Clustering K-Means and Weighted Product Method," *Proc. Atlantis Press Vol.165*, pp.18-22, July 2018.
- Raval, R. U., dan Chaita, J., Implementing and Improvisation of K-means Clustering Algorithm, *International Journal of Computer Science and Mobile Computing*, vol 5, hal 191-203, 2016.
- Zhang, C., dan Fang, Z., An improved k-means clustering algorithm, *Journal of Information and Computational Science*, Vol 10, hal 193-199, 2013,
- A. K. Wardhani, "Penerapan Algoritma Partitioning Around Medoids Untuk Menentukan Kelompok Penyakit Pasien(Studi Kasus : Puskesmas Kajen Pekalongan), *Jurnal Kilat*, Vol. 6, no 1, pp. 6-10, 2017.
- Zheng Y., dan Qian Li., Density k-means: a new algorithm for centers initialization for k-means, *IEEE*, hal 958-961, 2015.
- X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. Mclachlan, A. Ng, B. Liu, P.S. Yu, Z.Z. Michael, S. David, and J.H. Dan, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, 2008.