

Penerapan Metode K-Medoids Guna Pengelompokan Data Usaha Mikro, Kecil dan Menengah (UMKM) Bidang Kuliner Di Kota Yogyakarta

Utaminingsih Linarti¹, Anita Rahmawati², Anna Hendri Soleliza Jones³, Lisna Zahrotun⁴

^{2,3,4}Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

¹Teknik Industri, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Email: ¹utaminingsih.linarti@ie.uad.ac.id, ²anita.rahmawati@gmail.com, ³anna.henri@tif.uad.ac.id, ⁴lisna.zahrotun@tif.uad.ac.id

Abstrak—Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan kegiatan yang dikelola secara individu maupun lembaga dan berperan penting dalam pertumbuhan perekonomian negara. Dinas Perindustrian, Koperasi, dan Usaha Kecil dan Menengah Yogyakarta memiliki kendala mengenai pemilihan UMKM dalam memberikan layanan untuk meningkatkan UMKM di Kota Yogyakarta. Penelitian ini bertujuan untuk menghasilkan analisis dan akurasi data untuk mengelompokkan bidang Kuliner di kota Yogyakarta. Metode yang akan digunakan yaitu Algoritma K-Medoids. Metode ini merupakan algoritma pengelompokan untuk mempartisi data dan menetapkan objek representatif dengan kata lain medoids sebagai titik pusat atau centroid. Metode pengujian yang digunakan *silhouette coefficient* dengan langkah penelitian mulai dari load data, data cleaning, seleksi data, transformasi data, penerapan algoritma k-medoids, pengujian, dan representasi hasil. Penelitian menggunakan 1336 data dengan metode pengujian *silhouette coefficient* menghasilkan akurasi 0,60 dan menghasilkan pengelompokan data yang terbagi menjadi dua cluster. Cluster 0 memiliki data 1073 dan cluster 1 memiliki data 263. Pada Cluster 0 terdapat 70% memiliki pendapatan kurang dari 10 juta dan sebanyak 20% memiliki pendapatan 25 juta pertahun. Tujuan pemasaran dilakukan dalam wilayah DIY hingga luar Negeri dan rata-rata umur usaha 8,5 tahun. Pada cluster 1 terdapat 52% memiliki pendapatan kurang dari 10 juta pertahun dengan tujuan pemasaran dilakukan dalam wilayah DIY dan rata-rata umur usaha 33,6 tahun.

Kata Kunci: K-Medoids, Kuliner, UMKM, *Silhouette Coefficient*.

Abstract—*Micro, Small, and Medium Enterprises (MSMEs) are activities managed by individuals or institutions and play an important role in the country's economic growth. The Yogyakarta Office of Industry, Cooperatives, and Small and Medium Enterprises has constraints regarding the selection of MSMEs in providing services to improve MSMEs in the city of Yogyakarta. This research aims to produce data analysis and accuracy to classify the Culinary field in the city of Yogyakarta. The method that will be used is the K-Medoids Algorithm. This method is a clustering algorithm for partitioning data and determining representative objects in other words medoids as the centre point or centroid. The test method used is silhouette coefficient with research steps starting from data load, data cleaning, data selection, data transformation, application of the k-medoids algorithm, testing, and representation of results. Research using 1336 data with the silhouette coefficient testing method produced an accuracy of 0.60 and resulted in data grouping which was divided into two clusters. Cluster 0 has 1073 data and cluster 1 has 263 data. In Cluster 0, 70% have an income of less than 10 million and as much as 20% have an income of 25 million per year. Marketing objectives are carried out within the DIY region to abroad and the average age of the business is 8.5 years. In cluster 1, 52% have an income of less than 10 million per year with marketing destinations within the DIY region and an average business age of 33.6 years.*

Keywords: K-Medoids, Culinary, UMKM, *Silhouette Coefficient*.

1. PENDAHULUAN

Usaha Mikro, Kecil dan Menengah (UMKM) merupakan sebuah kegiatan yang dikelola secara individu maupun lembaga dan berperan penting dalam pertumbuhan perekonomian Negara dalam sebuah daerah. UMKM bertujuan untuk menumbuhkan dan mengembangkan usahanya dalam membangun perekonomian Negara. Dengan begitu UMKM ini menjadi salah satu penopang pertumbuhan ekonomi, dan berperan dalam mengatasi masalah pengangguran yang ada. Serta mudah diakses oleh pihak yang berkepentingan dan mewujudkan basis data UMKM melalui sistem pendataan yang terintegrasi dengan menggunakan teknologi informasi [1].

Teknologi informasi yang kini sedang ditingkatkan oleh pemerintah Kota Yogyakarta yaitu Jogja Smart Service (JSS), JSS merupakan sarana untuk promosi keunggulan UMKM dan informasi mengenai kegiatan pemerintah [2]. Pendataan UMKM sebelum masa pandemi covid-19 yang dilakukan oleh Dinas Perindustrian, Koperasi dan UMK pemerintah Kota Yogyakarta yaitu dengan cara terjun kelapangan untuk memastikan apakah UMKM yang sudah terdaftar benar adanya atau tidak. Tetapi saat masa pandemi covid-19 tidak dilakukan pendataan dengan terjun langsung ke lapangan dikarenakan adanya aturan pemerintah mengenai pembatasan *social distancing* untuk tidak berpergian, maka dari itu dilakukan melalui via whatsapp. Terdapat 32.441 data keseluruhan UMKM Kota Yogyakarta dari berbagai bidang yang belum terverifikasi. Sebanyak 2769 data sudah terverifikasi. Verifikasi data selanjutnya dilakukan berdasarkan beberapa bidang, salah satunya adalah bidang Kuliner. Terdapat 1336 data pada bidang kuliner. Dinas Perindustrian, Koperasi dan UMK pemerintah Kota Yogyakarta memiliki kendala yaitu pemilihan UMKM mana yang harus diberikan layanan pendampingan untuk meningkatkan UMKM bidang Kuliner di Kota Yogyakarta. Pemilihan tersebut dapat dilakukan dengan menggunakan data minning.

Kuliner merupakan produk olahan berupa makanan dan minuman, kuliner merupakan sebuah gaya hidup yang tidak dapat dipisahkan karena setiap orang membutuhkan makanan untuk kehidupan sehari-hari [3]. Kuliner di Kota

Yogyakarta menjadi salah satu daya tarik wisatawan yang berkunjung ke Kota Yogyakarta, selama masa pandemi covid-19 ribuan pelaku usaha mengalami keterpurukan sehingga masih membutuhkan dukungan agar bisa bangkit Kembali [4]. UMKM juga merupakan salah satu cara untuk mempercepat Pembangunan daerah [5]. Oleh sebab itu pemerintah kini sedang merencanakan untuk pendampingan UMKM salah satunya bidang kuliner.

Data Mining merupakan suatu proses dalam menganalisa suatu data yang sangat besar dengan perspektif yang berbeda sehingga menghasilkan suatu informasi yang berguna. Data Mining ini mampu membentuk dan mengenali pola baru dari kumpulan data dan informasi yang besar. Informasi dan pola baru yang terbentuk ini nantinya dapat digunakan sebagai solusi dari suatu permasalahan [6]. Salah satu teknik yang ada dalam *data mining* adalah *Clustering*. Salah satu penerapan clustering dalam UMKM adalah analisis pengaruh media social menggunakan *K-Means* [7]. Metode yang lainnya adalah *K-Medoids*, yang digunakan dalam bidang akademik [8] dan strategi marketing UMKM [5]. Pengelompokan UMKM bidang jasa juga pernah dilakukan sebelumnya menggunakan metode *Analytical Hierarchy Clustering* [9]. *Clustering* juga digunakan untuk persebaran UMKM [10].

Penelitian yang dilakukan oleh [8] dengan metode *K-Medoids* permasalahan yang terjadi yaitu pengelompokan masih dilakukan secara manual untuk hasil data penjualan, oleh karena itu dilakukannya pengelompokan data penjualan secara akurat. Dari pengelompokan data yang telah dilakukan pada tahun 2018 dan 2019 diperoleh dengan 3 hasil cluster yang terbagi kedalam cluster tinggi terdapat 2 items, cluster sedang terdapat 1 items dan cluster rendah terdapat 9 items. Metode *K-Medoids* sangat cocok digunakan untuk proses pengelompokan data-data. Metode *K-Medoids* merupakan algoritma pengelompokan untuk mempartisi data dan menetapkan objek sebagai titik pusat atau centroid [9]. Tujuan penelitian ini adalah mengimplementasikan penerapan metode *K-Medoids* dalam mengelompokan data UMKM bidang kuliner, yang mana hasil kelompok kelompok UMKM dapat memberikan gambaran UKM mana sajay yang masih membutuhkan pendampingan dan UMKM mana saja yang sudah berjalan dengan baik. Diharapkan dengan adanya sistem ini dapat mempermudah Dinas untuk memberikan layanan dalam meningkatkan UMKM melalui pendampingan-pendampingan yang dilakukan oleh Dinas kepada UMKM.

2. METODOLOGI PENELITIAN

2.1 UMKM

Usaha Mikro, Kecil dan Menengah (UMKM) merupakan sebuah kegiatan yang dikelola secara individu maupun lembaga dan berperan penting dalam pertumbuhan perekonomian Negara dalam sebuah daerah. UMKM bertujuan untuk menumbuhkan dan mengembangkan usahanya dalam membangun perekonomian Negara. UMKM ini menjadi salah satu penopang pertumbuhan ekonomi, dan berperan dalam mengatasi masalah pengangguran yang ada.

2.2 Algoritma K-Medoids

Metode yang digunakan yaitu metode *K-Medoids*. Algoritma *K-Medoids* merupakan algoritma pengelompokan untuk mempartisi data menetapkan objek representatif dengan kata lain medoids sebagai titik pusat atau centroid. Partisi algoritma *K-Medoids* untuk meminimalkan jumlah perbedaan antara setiap objek dalam data dan medoids terdekat [9, 11]. Algoritma *K-Medoids* yang cukup efisien untuk mengatasi data outlier dan tiap data akan dihitung kedekatan jaraknya dengan data lain pada tiap data. data pusat yang menjadi cluster merupakan hasil data yang diambil secara acak, data apapun dapat menjadi data pusat pada sebuah cluster yang sesuai dengan ketentuan dari Algoritma *K-Medoids*. Adapun tahap-tahap dari Algoritma *K-Medoids*, sebagai berikut :

- Menginisiasi pusat cluster sebanyak *K* (jumlah cluster)
- Kelompokkan setiap data pada cluster terdekat dengan cara menghitung jarak menggunakan Euclidean Distance pada masing-masing medoids terhadap suatu objek antar data yang dinyatakan pada persamaan (1) :

$$d(x, y) = \sqrt{(\sum_{i=1}^n (x_{(i)} - y_{(i)}))^2}; 1,2,3 \dots n$$

pers. (1)

Keterangan :

- $x_{(i)}$ = data *i* yang pertama
- $y_{(i)}$ = data *i* yang kedua
- n* = banyaknya data
- Memilih secara tidak teratur pada objek setiap cluster sebagai calon medoids baru.
- Hitung jarak Euclidean untuk setiap objek terhadap tiap-tiap cluster dengan calon medoids baru.
- Hitung total simpangan (*S*) dari fungsi objektif dengan cara hitung nilai total jarak baru – total jarak lama. Jika $S < 0$, lalu objek tersebut ditukar menggunakan data cluster untuk menghasilkan kumpulan cluster baru sebagai medoids.
- Ulangi langkah (c-e) hingga tidak ada lagi perubahan medoids kemudian dihasilkan sekelompok cluster yang sesuai.

2.3 Pengujian

Metode *silhouette coefficient* digunakan pada pengujian penelitian ini, metode ini menghitung jarak kedekatan dan jarak keterpisahan tiap data atau objek dalam cluster [10]. Langkah-langkah pada *silhouette coefficient* [12,13], sebagai berikut:

- a. Hitung rata-rata jarak terhadap suatu dokumen ke- i dengan tiap dokumen lainnya yang terdapat pada satu cluster yang sama menggunakan persamaan (2).

$$a_i = \frac{1}{n_p - 1} \sum_{r=1}^{n_p - 1} d_{i,r}, r \neq i \quad \text{pers. (2)}$$

- b. Hitung rata-rata jarak terhadap dokumen ke- i dengan tiap dokumen yang terdapat pada satu cluster yang berbeda menggunakan persamaan (3), lalu mengambil nilai terkecil dengan persamaan (4).

$$b_i = \min \{d_i(p)\}, r \neq i \quad \text{pers. (3)}$$

dengan rumus jarak suatu dokumen ke- i dengan tiap data pada cluster yang lainnya.

$$d_i(p) = \frac{1}{n_p} \sum_{r=1}^{n_p} d_{i,r} \quad \text{pers. (4)}$$

- c. Menghitung nilai *silhouette coefficient* dapat dilihat pada persamaan (5)

$$SC_1(i) = \frac{b_i - a}{\max\{a_i, b_i\}}, i = 1, 2, \dots, n \quad \text{pers. (5)}$$

Keterangan :

$SC_1(i)$ = nilai silhouette coefficient

a_i = jarak rata-rata i dengan seluruh objek pada cluster yang sama.

b_i = jarak rata-rata i dengan seluruh objek pada cluster lain.

2.4 Tahapan Penelitian

1. Pengumpulan Data
Data tersebut diperoleh dari pendataan di Dinas Perindustrian, Perdagangan dan Koperasi Kota Yogyakarta.
2. Seleksi Data
Tahapan pemilihan data relevan yang dibutuhkan untuk keperluan analisis. Dalam pemilihan data Pendidikan, kegiatan usaha, tujuan pemasaran, status kepemilikan tanah dan bangunan, sarana media elektronik, modal bantuan pemerintah, pinjaman kredit usaha rakyat, omset pertahun, kepemilikan asuransi kesehatan, tenaga kerja laki-laki, tenaga kerja perempuan dan tanggal pendirian usaha.
3. Transformasi Data
Proses untuk melakukan perubahan dan pengkategorian data berupa *text* kedalam format biner 0 dan 1 dengan menggunakan teknik *One Hot Encoding*, kegunaan teknik ini untuk memudahkan pemrosesan data yang akan dilakukan [14, 15, 16].
4. K-Medoids Pengelompokan
Tahapan ini merupakan proses pengelompokan (*clustering*) data yang telah ditentukan pada tahap transformasi dengan menggunakan algoritma *K-Modes* dan *Cohesion and Separation* serta penggunaan nilai k sebesar 2, 3, 4, 5, ..., 10 agar kamu dapat mengetahui nilai k . terbaik untuk menentukan pola kecelakaan.
5. Penentuan Optimal K
Tahapan penentuan K Optimal dilakukan untuk melihat nilai K terbaik yang diperoleh dari hasil clustering k mode. Penentuan Optimal K dilakukan dengan melakukan percobaan sebanyak 10 kali
6. Analisis Pola
Analisis pola merupakan informasi dari penelitian yang telah dilakukan yang akan memberikan hasil kepada pengguna dari hasil clustering, pola yang dihasilkan pada penelitian ini berupa hasil *centroid* dan anggota cluster.

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data UMKM bidang kuliner di Kota Yogyakarta yang berjumlah 1336 data. Data ini merupakan data UMKM tahun 2021-2022 dengan variabel, yaitu: (1) Pendidikan, (2) Kegiatan usaha, (3) Tujuan pemasaran, (4) Status kepemilikan tanah/anggunan, (5) Sarana media elektronik, (6) Modal bantuan pemerintah, (7) Pinjaman kredit usaha rakyat, (8) Omset pertahun, (9) Kepemilikan asuransi kesehatan, (10) Tenaga kerja laki-laki, (11) Tenaga kerja perempuan dan (12) Lama pendirian usaha. Penentuan variabel ini merupakan *state of the art* dalam penelitian ini berdasarkan *brainstroming* dengan Dinas Perindustrian, Koperasi, dan Usaha Kecil dan Menengah Yogyakarta. Variabel yang banyak digunakan dalam cluster UMKM adalah omset penjualan, aset dan jumlah tenaga

kerja [17,18]. Selain itu, penelitian ini berfokus pada salah satu bidang UMKM, yaitu bidang kuliner. Berdasarkan [19], cluster UMKM ditentukan berdasarkan variabel kecamatan.

Setelah dilakukan seleksi maka akan dilakukan proses transformasi data. Transformasi data dalam penelitian ini adalah mengubah variabel yang digunakan menjadi variabel yang dapat dikelompokkan menggunakan metode K-Medoid. Berbeda dengan transformasi yang dilakukan sebelumnya yaitu (1) Memahami literasi digital; (2) Merencanakan sumber daya yang dibutuhkan dan menganalisis kebutuhan konsumen digital; (3) Mendesain ulang; (4) Mempersiapkan anggaran; (5) Melakukan kemitraan; dan (6) implementasi . Transformasi data bentuk kualitatif dalam penelitian ini ditunjukkan pada Tabel 1 1.

Tabel 1. Hasil Transformasi Data

No	Variabel	Hasil Variabel dan One-Hot Encoding
1	Pendidikan Terakhir	-, 0, D1, D2, D3, D4, S1, S2, S3, SD, SMA, SMK, SMP
2	Kegiatan Usaha	Penjualan, Produksi
3	Tujuan Pemasaran	Dalam wilayah DIY, Dalam wilayah Kota Yogyakarta, Dalam Wilayah pulau Jawa, Dalam wilayah luar pulau Jawa (tersebar), Luar Negeri (ekspor), Luar Negeri (jenis pengiriman biasa)
4	Status Kepemilikan Tanah/Bangunan	Lainnya, Magersari (adat), Milik sendiri, Sewa
5	Sarana Media Elektronik	-, Facebook, Gojek, Grab, Instagram, Shopee, Tokopedia, Twitter, WhatsApp, Lainnya
6	Modal Bantuan Pemerintah	-, Pemda DIY, Pemerintah Pusat, Pemkot Yogyakarta
7	Pinjaman Kredit Usaha Rakyat	-, Bank, Koperasi, Lainnya, Pemerintah
8	Omset per-Tahun	Kurang dari 10 juta, 10 juta s/d 25 juta, 40 juta s/d 55 juta, 55 juta s/d 70 juta, 70 juta s/d 85 juta, 85 juta s/d 100 juta, 100 juta s/d 120 juta, 120 juta s/d 150 juta, Lebih dari 150 juta
9	Kepemilikan Asuransi Kesehatan	-, Asuransi Swasta, BPJS

Adapun implementasi aplikasi ini adalah sebagai berikut:

3.1 Load Data

Dataset yang akan digunakan berjumlah 1336 data UMKM kuliner akan diload terlebih dahulu sebelum dilakukan proses preprocessing. Load data dapat dilihat pada Gambar 1.

Gambar 1 Load Data

Ref. OSS	Nomor Induk Kependudukan	Nama Lengkap	Tanggal Lahir	Usia	Jenis Kelamin	Pendidikan Terakhir	No. Telepon	e-Mail	Provinsi	Status Kepemilikan Tanah/Bangunan	
0	17678	3.471015e+15	ALYIAH	11 Maret 1961	41	P	SMA	89620770486	nasywap189@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1	12185	3.471046e+15	A. ENDANG KASMIATI	15 Februari 1969	53	P	SMA	85718338652	andreaenaendang45@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
2	6808	3.471014e+15	A. RIKA NURHARYANTI TRISNO	02 Maret 1977	45	P	0	81226224778	rintrika@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
3	8723	3.471131e+15	AAN DWI YANTO	12 April 1981	40	L	SMA	89630991456	s@g.com	DAERAH ISTIMEWA YOGYAKARTA	Sewa
4	4085	3.403131e+15	ABDUL AZIS	-	-	L	SMP	85743612564	aziza9647@gmail.com	DI. YOGYAKARTA	Lainnya
...
1331	3956	3.471030e+15	YUSUP	31 Oktober 1975	46	L	-	89504355811	arispuji1974@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1332	9534	3.217092e+15	YUSUP JAMHUR	18 Juni 1978	43	L	SMA	89504946740	pangestinana80@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
1333	11324	3.471015e+15	YUYUN AFRIANTI	10 Oktober 1974	47	P	S1	85228442475	ananda.jukti@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
1334	13178	3.471071e+15	YUYUN TRIYANTO	06 Juni 1985	36	L	SMK	88980953428	faraicha06@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1335	15213	3.471145e+15	Z U H A N A	12 Oktober 1957	64	P	SD	8995138283	zuhananyonya@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya

1336 rows x 34 columns

3.2 Data Cleaning

Proses cleaning dilakukan untuk membersihkan atau menghilangkan value yang tidak diperlukan dalam proses penelitian ini, dari 1336 data yang digunakan untuk proses cleaning tidak ditemukan data yang tidak memiliki nilai, sehingga datanya tidak berkurang. Gambar 2 merupakan gambaran tabel proses data *cleaning*.

Gambar 2 Data Cleaning

Ref. OSS	Nomor Induk Kependudukan	Nama Lengkap	Tanggal Lahir	Usia	Jenis Kelamin	Pendidikan Terakhir	No. Telepon	e-Mail	Provinsi	Status Kepemilikan Tanah/Bangunan	
0	17676	3.471015e+15	ALYAH	11 Maret 1981	41	P	SMA	89620770486	nasywap189@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1	12185	3.471046e+15	A. ENDANG KASMIATI	15 Februari 1969	53	P	SMA	85718338652	andreaenaendang45@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
2	6808	3.471014e+15	A R IKA NURHARYANTI TRISNO	02 Maret 1977	45	P	0	81226224778	rintrika@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
3	8723	3.471131e+15	AAN DWI YANTO	12 April 1981	40	L	SMA	89630991456	s@g.com	DAERAH ISTIMEWA YOGYAKARTA	Sewa
4	4085	3.403131e+15	ABDUL AZIS	-	-	L	SMP	85743612564	aziza9647@gmail.com	DI YOGYAKARTA	Lainnya
...
1331	3956	3.471030e+15	YUSUP	31 Oktober 1975	46	L	-	89504355811	arispuji1974@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1332	9534	3.217092e+15	YUSUP JAMHUR	18 Juni 1978	43	L	SMA	89504946740	pangestinan80@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
1333	11324	3.471015e+15	YUYUN AFRIANTI	10 Oktober 1974	47	P	S1	85228442475	ananda.jukil@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Milik sendiri
1334	13178	3.471071e+15	YUYUN TRIYANTO	06 Juni 1985	36	L	SMK	88980953428	faraicha06@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya
1335	16213	3.471145e+15	Z U H A N A	12 Oktober 1987	64	P	SD	8995138283	zuhananyonya@gmail.com	DAERAH ISTIMEWA YOGYAKARTA	Lainnya

1336 rows x 34 columns

3.3 Seleksi Data

Seleksi data digunakan untuk menentukan data yang akan digunakan untuk pemrosesan data. Penelitian ini dilakukan seleksi data untuk menyeleksi variabel yang akan digunakan. Pada awalnya data ini memiliki 34 variabel didalamnya, kemudian dilakukan seleksi data menjadi 12 variabel yang digunakan. Seleksi data yang dilakukan secara lengkap dapat dilihat pada Gambar 3.

Gambar 3 Seleksi Data

Pendidikan	Tgl_Pendirian_Usaha	Kegiatan_Usaha	Tujuan_Pemasaran	Status_Kepemilikan_Tanah	Sarana_Media_Elektronik	Modal_Bantuan_Pemerintah	
0	SMA	24 Juli 2012	Penjualan, Produksi	Dalam wilayah DIY	Lainnya	WhatsApp, Facebook	-
1	SMA	16 Februari 2016	Penjualan	Dalam wilayah Kota Yogyakarta, Dalam wilayah D...	Milik sendiri	WhatsApp, Instagram	Pemkot Yogyakarta
2	0	07 Januari 2022	Penjualan, Produksi	Dalam wilayah Kota Yogyakarta	Lainnya	WhatsApp	Pemkot Yogyakarta
3	SMA	22 Juni 2020	Penjualan	Dalam wilayah Kota Yogyakarta	Sewa	WhatsApp	Pemkot Yogyakarta
4	SMP	04 Februari 2005	Penjualan	Dalam wilayah Kota Yogyakarta	Lainnya	Lainnya	Pemerintah Pusat
...
1331	-	20 Agustus 1990	Penjualan	Dalam wilayah Kota Yogyakarta	Lainnya	WhatsApp	-
1332	SMA	26 Oktober 2011	Penjualan, Produksi	Dalam wilayah Kota Yogyakarta	Milik sendiri	WhatsApp	-
1333	S1	09 September 2020	Penjualan, Produksi	Dalam wilayah DIY	Milik sendiri	WhatsApp, Facebook, Twitter, Instagram	-
1334	SMK	07 September 2020	Penjualan	Dalam wilayah Kota Yogyakarta	Lainnya	-	-
1335	SD	17 Oktober 2018	Penjualan, Produksi	Dalam wilayah Kota Yogyakarta	Lainnya	WhatsApp	-

1336 rows x 12 columns

3.4 Transformasi Data

Penelitian ini menggunakan sebuah metode untuk melakukan transformasi data yaitu one hot encoding untuk menginisialisasi format data kedalam binary agar dapat dilakukan pemrosesan data. Data yang digunakan pada proses transformasi yaitu data yang sudah diseleksi dapat dilihat pada Gambar 4.

Gambar 4 Transformasi Data

	AKADEMI/DIPLOMA III/SARJANA MUDA	BELUM TAMAT SD/SEDERAJAT	D1	D2	D3	D4	DIPLOMA IV/STRATA I	S1	...	juta s/d 85 juta	juta s/d 100 juta	Kurang dari 10 juta	Lebih dari 150 juta	Asuransi Swasta	BPJS	Tenaga_Kerja_L	Tenaga
0	0	0	0	0	0	0	0	0	...	0	0	1	0	1	0	0	0
1	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	1	0
2	0	1	0	0	0	0	0	0	...	0	0	0	0	1	0	0	1
3	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	1	1
4	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	1	0
...
1331	1	0	0	0	0	0	0	0	...	0	0	1	0	1	0	0	2
1332	0	0	0	0	0	0	0	0	...	0	0	0	0	1	0	0	1
1333	0	0	0	0	0	0	0	1	...	0	0	1	0	0	0	1	0
1334	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	1	1
1335	0	0	0	0	0	0	0	0	...	0	0	1	0	0	0	1	0

1336 rows x 64 columns

3.5 Algoritma K-Medoids

Pada tahap ini seluruh data akan diukur menggunakan *Euclidean distance* untuk mencari jarak kedekatan tiap data. Pengolahan data menggunakan algoritma K-Medoids dapat dilihat pada Gambar 5.

Gambar 5 Algoritma K-Medoids

id	1	2	3	4	5	6	7	8	9	10
1	0	31	114	79	63	75	94	93	491	78
2	31	0	53	30	138	32	185	36	690	31
3	114	53	0	15	305	32.41421	378	21	1036.414	19.4142136
4	79	30	15	0	236	34	309	16	914	15
5	63	138	305	236	0	232	23	270	238	239
6	75	32	32.41421	34	232	0	285	36	868	34.4142136
7	94	185	378	309	23	285	0	333	179	300
8	93	36	21	16	270	36	333	0	970	11
9	491	690	1036.414	914	238	868	179	970	0	908.414214
10	78	31	19.41421	15	239	34.41421	300	11	908.4142	0

3.6 Hasil Clustering

Hasil dari perhitungan menggunakan algoritma *K-Medoids* akan mendapatkan nilai akhir berupa jumlah *cluster*. Penentuan *cluster* berupa data-data yang sudah dikelompokkan berdasarkan tingkat kemiripan antar data. *Cluster* yang terbentuk berdasarkan variabel yang telah ditetapkan melalui tahapan pengujian sebagai proses validasi.

Tabel 6 Hasil Clustering

dia_Elektronik	Modal_Bantuan_Pemerintah	Pinjaman	Omset_Pertahun	Kepemilikan_Asuransi_Kesehatan	Tenaga_Kerja_L	Tenaga_Kerja_P	Umur_Usaha	Cluster
App, Facebook	-	-	Kurang dari 10 juta	-	0	0	10.0	0
App, Instagram	Pemkot Yogyakarta	-	Kurang dari 10 juta	-	BPJS	0	2	6.0
WhatsApp	Pemkot Yogyakarta	Bank, Pemerintah	10 juta s/d 25 juta	-	1	1	0.0	0
WhatsApp	Pemkot Yogyakarta	Lainnya	10 juta s/d 25 juta	-	BPJS	1	0	2.0
Lainnya	Pemerintah Pusat	Lainnya	Kurang dari 10 juta	-	BPJS	0	0	17.0
...
WhatsApp	-	-	Kurang dari 10 juta	-	2	0	32.0	1
WhatsApp	-	-	10 juta s/d 25 juta	-	1	0	11.0	1
App, Facebook, twitter, Instagram	-	-	Kurang dari 10 juta	-	BPJS	0	0	2.0
-	-	Lainnya	Kurang dari 10 juta	-	BPJS	1	0	2.0
WhatsApp	-	-	Kurang dari 10 juta	-	BPJS	0	2	4.0

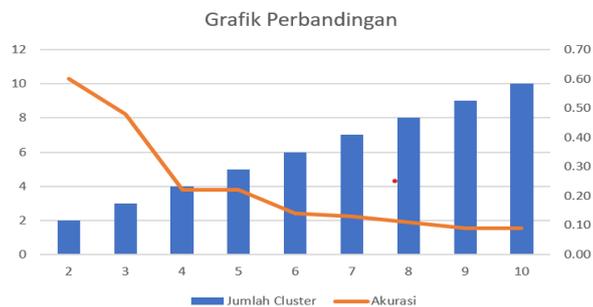
3.7 Pengujian

Data yang telah dilakukan pemrosesan menggunakan algoritma K-Medoids selanjutnya akan dilakukan tahapan pengujian. Pengujian yang menghasilkan proses yang efisien dilakukan dengan *Silhouette Coefficient* (SC) [20]. Tabel 2 merupakan hasil pengujian yang telah dilakukan dilihat dari kriteria akurasi SC dan waktu proses *running*. Berdasarkan Tabel 2 jumlah kluster terbaik adalah 2, dimana nilai akurasi SC memiliki nilai paling tinggi sebesar 0,60 dan memiliki waktu proses tercepat sebesar 07.71 detik.

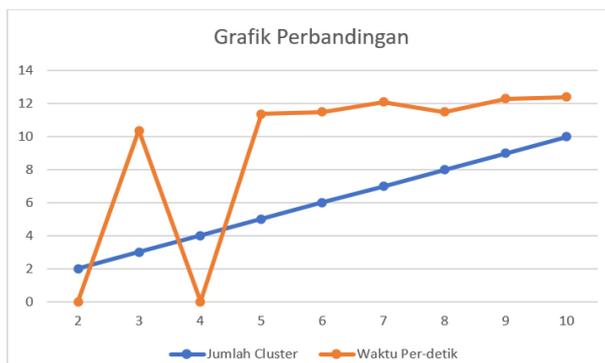
Tabel 2. Hasil Pengujian

No	Jumlah Cluster	Akurasi	Waktu (detik)
1	2	0.60	07,71
2	3	0.48	10,37
3	4	0.22	09,88
4	5	0.22	11,37
5	6	0.14	11,49
6	7	0.13	12,10
7	8	0.11	11,50
8	9	0.09	12,30
9	10	0.09	12,40

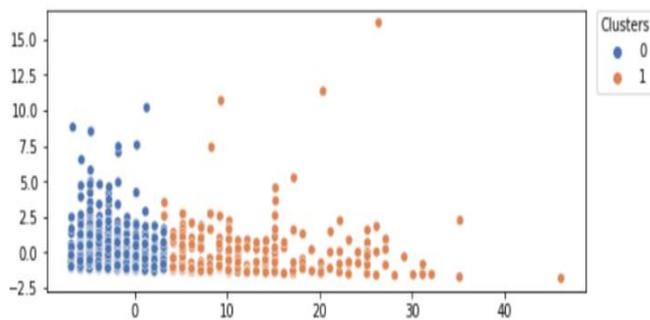
Grafik perbandingan jumlah cluster dan akurasi ditunjukkan dalam Gambar 7. Grafik perbandingan jumlah cluster dengan waktu eksekusi ditunjukkan pada Gambar 8 dan sebaran data hasil cluster ditunjukkan pada Gambar 9.



Gambar 7 Grafik perbandingan jumlah cluster dan akurasi



Gambar 8 Grafik perbandingan jumlah cluster dan waktu



Gambar 9 Sebaran hasil Clustering

Berdasarkan Gambar 7 dan Gambar 8 dapat dilihat secara visual bahwa semakin besar jumlah *cluster* pola pengujian yang terbentuk adalah linier, dimana SC akan semakin kecil dan waktu proses semakin lama. Jumlah *cluster* dari 2 ke 3 mengalami penurunan angka pengujian SC dan waktu proses. Namun jumlah *cluster* dari 3 ke 4 terdapat perubahan pola linier yang terjadi, sedangkan pada jumlah *cluster* 5 ke atas semakin kecil nilai SC dan semakin besar waktu proses pengolahan data. Gambar 9 nampak bahwa jumlah UMKM pada *cluster* 1 memiliki persentase anggota UMKM yang lebih banyak dibandingkan anggota *cluster* 0. Selain itu, nampak bahwa ada anggota dalam masing-masing *cluster* yang menunjukkan kondisi outlier kemiripan dalam klasifikasi anggota. Meskipun demikian tampak bahwa anggota pada *cluster* 0 lebih menunjukkan kemiripan antar anggota. Hal tersebut menarik terkait dengan pengujian SC yang mengalami perubahan pola linier pada jumlah *cluster* 4 dan sebaran kemiripan anggota berdasarkan variabel yang telah ditetapkan.

3.8 Representasi Pengetahuan

1) Cluster 0

Hasil representasi pengetahuan yang didapat pada *cluster* 0 memiliki 1073 data UMKM bidang kuliner, pendapatan pada *cluster* 0 sekitar 70% kurang dari 10 juta dan 20% 25 juta dengan Pendidikan 60% SMA, SMK dan S1, serta kegiatan usaha yang dilakukan yaitu penjualan, penjualan dan produksi. Tujuan pemasaran dilakukan dalam wilayah DIY, Kota Jogja, luar pulau Jawa bahkan sampai ke luar Negeri, *cluster* 0 ada yang mendapatkan bantuan dari pemerintah sekitar 30% dan 70% tidak mendapatkan bantuan dari pemerintah dan cenderung tidak melakukan pinjaman. Serta kepemilikan bangunan usaha 55% milik sendiri dan sewa, rata-rata tenaga kerja yang dimiliki 1 hingga 2 orang cenderung memiliki asuransi berupa BPJS. Rata-rata umur usaha selama 8.5 tahun

2) Cluster 1

Hasil representasi pengetahuan yang didapat pada *cluster* 1 memiliki 263 data UMKM bidang kuliner, pendapatan pada *cluster* 1 sekitar 52% kurang dari 10 juta dengan Pendidikan 62% SMA, SMK dan SMP, serta kegiatan usaha yang dilakukan yaitu penjualan dan penjualan, produksi. Tujuan pemasaran dilakukan cenderung dalam wilayah DIY dan Kota Jogja, *cluster* 1 banyak yang tidak mendapatkan bantuan dari pemerintah sekitar 70% dan 60% tidak melakukan pinjaman. Serta kepemilikan bangunan usaha 53% milik sendiri dan sewa, rata-rata tenaga kerja yang dimiliki 3 hingga 4 orang cenderung memiliki asuransi berupa BPJS. Rata-rata umur usaha selama 33.6 tahun.

4. KESIMPULAN

Variabel yang digunakan dalam pembentukan *cluster* penelitian ini merupakan hasil *brainstroming* dengan stakeholder terkait dan dapat digunakan sebagai acuan pada penelitian lain khususnya UMKM. Berdasarkan analisis hasil dapat disimpulkan bahwa penelitian ini menghasilkan aplikasi pengelompokan data UMKM bidang kuliner menggunakan metode *K-Medoids*. Berdasarkan pengujian *silhouette coefficient* diperoleh akurasi dan waktu proses terbaik dengan jumlah 2 *cluster*. Dari 2 *cluster* didapatkan rekomendasi untuk memberikan pendampingan kiat pemasaran dan kapasitas produksi pada UMKM di cluster 0 yaitu sejumlah 752 dari 1073 UMKM yang masih memiliki omset di bawah 10 juta pertahun. Sedangkan untuk cluster 1 diberikan pendampingan dalam teknik pemasaran dan pengembangan kapasitas produksi karena masih terdapat 137 dari 263 UMKM pemasarannya masih dalam wilayah DIY dan omsetnya masih di bawah 10 juta. Perlu dilakukan kajian lebih dalam oleh Dinas Perindustrian, Koperasi dan Usaha Kecil Menengah terkait masih adanya 70% anggota pada *cluster* 0 yang tidak mendapatkan bantuan dari pemerintah. Pemetaan anggota untuk masing-masing *cluster* perlu dilakukan, sehingga akan memudahkan analisis rekomendasi pendampingan untuk meningkatkan efektivitas dan produktivitas UMKM bidang kuliner yang ada di Kota Yogyakarta.

Pada penelitian ini masih terdapat kekurangan, maka perlu adanya pengembangan untuk ke depannya untuk memperoleh hasil akurasi yang lebih baik. Saran yang diberikan yaitu menggunakan metode lain untuk perbandingan pengelompokan data agar mengetahui hasil *cluster* yang baik.

UCAPAN TERIMAKASIH

Pada penelitian ini kami mengucapkan terimakasih kepada LPPM UAD atas pendanaan yang diberikan dalam mendukung terlaksananya penelitian ini.

REFERENCES

- [1] M. Hasan, A. Dzakiyyah, D. A. Kumalasari, N. Safira, and S. N. Aini, "Transformasi Digital UMKM Sektor Kuliner Di Kelurahan Jatinegara, Jakarta Timur," *J. Bisnis dan Kewirausahaan*, vol. 17, no. 2, pp. 135–150, 2021, doi: 10.31940/jbk.v17i2.2529.

- [2] L. Putri, "Pemkot Yogyakarta Luncurkan Jogja Smart Service (JSS) Versi 3," 2021.
- [3] Y. Tresnawati and K. Prasetyo, "Pemanfaatan Digital Marketing Bagi Usaha Mikro Kecil dan Menengah Bisnis Kuliner," *J. New Media Commun.*, vol. 1, no. 1, pp. 43–57, 2022, doi: 10.55985/jnmc.v1i1.5.
- [4] F. N. Aqilla, "Kajian UMKM Dalam Rangka Pemulihan Ekonomi Daerah Kabupaten Bogor Tahun 2021," *Bina J. Pembang. Drh.*, vol. 1, no. 1, pp. 101–125, 2021.
- [5] D. U. Iswavigra, L. E. Zen, and H. Hanim, "Marketing Strategy UMKM Dengan CRISP-DM Clustering & Promotion Mix Menggunakan Metode K-Medoids," *J. Inf. dan Teknol.*, vol. 5, pp. 45–54, 2023, doi: 10.37034/jidt.v5i1.260.
- [6] J. Han, M. Kamber, and J. Pei, *Introduction*. 2012. doi: 10.1016/b978-0-12-381479-1.00001-0.
- [7] W. Aristika and W. J. Hartono, "Penerapan Clustering K-Menas Untuk Menentukan Pengaruh Media sosial Facebook Terhadap Usaha Mikro, KECil dan Menengah (UMKM) Di Kecamatan Pekanbaru Kota," *J. Ilmu Komput. dan Bisnis*, vol. 11, no. 1, pp. 2389–2395, 2020.
- [8] H. Kurnia, L. Zahrotun, and U. Linarti, "Pengelompokan Mahasiswa Berdasarkan Data Akademik Sebelum Kuliah Dan Masa Studi Menggunakan K-Medoids," *Infotekjar J. Nas*, vol. 2, 2021, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/3243>
- [9] L. Zahrotun, S. H. Nugroho, U. Linarti, and A. H. S. Jones, "Analisis Persebaran UMKM Bidang Jasa Menggunakan Metode AHC Complete Linkage," *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 2, pp. 255–265, 2023.
- [10] D. Remawati, D. J. Aji Putra, and T. Irawati, "Metode K-Means Untuk Pemetaan Persebaran Usaha Mikro Kecil Dan Menengah," *J. Teknol. Inf. dan Komun.*, vol. 9, no. 2, p. 39, 2021, doi: 10.30646/tikomsin.v9i2.574.
- [11] N.K. Kaur, U. Kaur, and D. Singh, "K-Medoid Clustering Algorithm - A Review", *Int. J. Comput. Appl. Technol.*, vol. 1, pp. 42-45, 2014,
- [12] R. U. B. Barus, I. Gunawan, B. E. Damanik, I. Parlina, and W. Saputra, "Pengelompokan Data Penjualan Mie Berdasarkan Bulan Dengan Menggunakan Algoritma K-Medoids," *J. Ilmu Komput. dan Inform.*, vol. 1, no. 2, pp. 141–156, 2021, doi: 10.54082/jiki.15.
- [13] J. Han, M. Kamber, and J. Pei, "Third Edition : Data Mining Concepts and Techniques," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2012.
- [14] R. N. Ibrahim and M. N. Hayati, "Penerapan Algoritma K-Medoids pada Pengelompokan Wilayah Desa atau Kelurahan di Kabupaten Kutai Kartanegara," *J. EKSPONENSIAL*, vol. 11, pp. 153–158, 2020.
- [15] K. E. Y. Features, "OPC Toolbox 2".
- [16] I. W. Saputro and B. W. Sari, "Uji Performa Algoritma Naïve Bayes untuk Prediksi Masa Studi Mahasiswa," *Creat. Inf. Technol. J.*, vol. 6, no. 1, p. 1, 2020, doi: 10.24076/citec.2019v6i1.178.
- [17] T. Zulyanti and Noeryanti, "Perbandingan Pengelompokan USaha Mikro Kecil dan Menengah Kabupaten Klaten Tahun 2019 dengan Metode K-Means dan ClusteringLarge Application", *J. Statistika Industri dan Komputasi*, vo. 7, no. 1, pp. 46-59, 2022
- [18] D.J.A. Putra, D Remawati and T. Irawati, "Metode K-Means untuk Pemetaan Persebaran Usaha Mikro Kecil dan Menengah", *J.TIKomSI*, vol. 9, no. 2, pp. 39-46, 2021
- [19] R. Wahyusari and W. Setia, "Perbandingan Algoritma K-Means dan Algoritma K-Medoids untuk Pengelompokan UMKM di Kebumen", *Proceeding Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat Bidang Ilmu Komputer Universitas PGRI Madiun*, vol. 2, pp. 73-79, 2023
- [20] M.B. Al-Zoubi and M.A. Rawi, "An Efficient Approach for Computing Silhouette Coefficients", *J. Comput. Sci.*, vol. 4, no. 3, pp.252-255, 2008