

Marketing Strategy UMKM dengan CRISP-DM Clustering dan Promotion Mix Menggunakan Metode K-Medoids

Dwi Utari Iswavigra^{1□}, Lova Endriani Zen², Okfalisa³, Hafizah Hanim⁴

^{1,2}Institut Teknologi Sains dan Kesehatan Sugeng Hartono, Indonesia

³Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

⁴Universitas Andalas, Indonesia

dwi.utari.iswavigra1997@gmail.com

Abstrak

Perkembangan UMKM merupakan suatu hal yang berpengaruh sangat besar terhadap perekonomian di Indonesia. Pasar UMKM lokal bahkan mampu bersaing secara global yang ditunjukkan dengan banyaknya permintaan dari luar negeri yang masuk. Meskipun begitu, UMKM di Indonesia pernah mengalami fluktuasi akibat krisis ekonomi. Fluktuasi yang terjadi menunjukkan penurunan jumlah UMKM sebanyak 0,003%. Direktur Eksekutif *Institute For Economic* (Indef) mengatakan bahwa fluktuasi yang terjadi pada UMKM, mengakibatkan perkembangan ekonomi menjadi tidak stabil pada kuartal II. Dari masalah ini, maka dilakukan proses Clustering CRISP-DM untuk mengolah data UMKM yang ada untuk menentukan strategi promosi yang tepat dalam mengembangkan UMKM berdasarkan jenis usaha yang dijalani, omset dan aset yang dimiliki. Penelitian ini dilakukan menggunakan algoritma K-Medoids dengan membentuk 3 cluster untuk proses pengolahan datanya. Data yang di olah sebanyak 71 data UMKM Se-Solo Raya, di mana pada cluster 1 sebanyak 25 UMKM yang di dominasi oleh jenis usaha di bidang fashion dengan rata-rata omset dan aset antara Rp.1.000.000 - Rp.5.000.000, cluster 2 sebanyak 39 UMKM yang di dominasi oleh jenis usaha di bidang kuliner dengan rata-rata omset dan aset antara Rp.1.000.000 - Rp.5.000.000 dan pada cluster 3 sebanyak 7 UMKM yang di dominasi oleh jenis usaha di bidang lain (diluar 6 jenis usaha lainnya) dengan rata-rata aset sebesar $\geq$ Rp.30.000.000 dan omset antara Rp.21.000.000 - Rp.25.000.000.

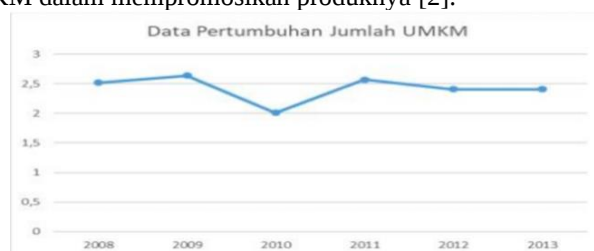
Kata Kunci: UMKM, Promotion Mix, K-Medoids, Clustering, CRISP-DM.

Jidt is licensed under a Creative Commons 4.0 International License.



1. Pendahuluan

UMKM merupakan sebutan bagi komunitas ekonomi mikro, kecil, menengah. Pertumbuhan UMKM sangat berpengaruh dalam perkembangan ekonomi di Indonesia. UMKM juga berpotensi untuk mempercepat pembangunan daerah. Pada tahun 1997, UMKM sempat mengalami krisis ekonomi, tetapi hal tersebut lantas tidak menurunkan pertumbuhan UMKM. Krisis tersebut bahkan dijadikan sebagai penopang dalam proses pemulihan ekonomi bangsa karena para pelaku UMKM mampu memberikan kontribusi yang besar terhadap penyerapan tenaga kerja [1]. Meskipun demikian, berdasarkan *update* terbaru dari Badan Pusat Statistik pada tanggal 22 Desember 2016, data menunjukkan bahwa terjadi fluktuasi persentase jumlah UMKM pada tahun 2008-2013 yang berakibat terjadinya penurunan sebesar 0,003% pada tahun 2013. Enny Sri Hartati selaku Direktur Eksekutif *Institute For Economic* (Indef) mengatakan bahwa fluktuasi yang terjadi pada UMKM, mengakibatkan perkembangan ekonomi menjadi tidak stabil pada kuartal II. Tentu saja hal ini juga memberikan dampak negatif pada pertumbuhan ekonomi negara. Pihak LPPI dan BI juga mengatakan bahwa ada banyak hal lainnya yang menyebabkan penurunan jumlah UMKM yaitu salah satunya belum terbentuknya strategi promosi yang baik dan tepat oleh para pelaku UMKM dalam mempromosikan produknya [2].



Gambar 1. Data Pertumbuhan Jumlah UMKM Tahun 2008-2013 (Sumber : BPS Nasional 2016)

Data Mining / CRISP-DM merupakan suatu proses dalam menganalisa suatu data yang sangat besar dengan perspektif yang berbeda sehingga menghasilkan suatu informasi yang berguna. Data Mining / CRISP-DM ini juga biasa dikenal sebagai Knowledge Discovery in Database (KDD) [3] [4]. Teknik 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.

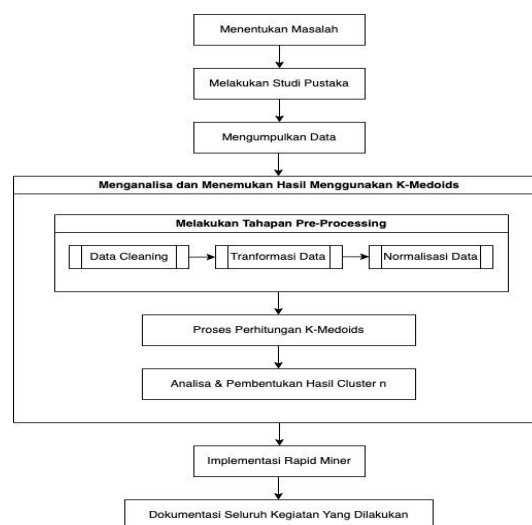
Data Mining / CRISP-DM memiliki beberapa metode salah satu diantaranya adalah *Clustering*. *Clustering* merupakan metode yang digunakan untuk menemukan struktur *cluster* yang memiliki kesamaan dan mengelompokkannya menjadi 1 *cluster* [5]. Salah satu bentuk algoritma yang biasa digunakan dalam proses *Clustering* ini adalah K-Medoids. Algoritma K-Medoids merupakan algoritma *Clustering* non-hierarkis yang diturunkan dari algoritma K-Means. Algoritma K-Medoids ini dibentuk sebagai solusi dalam memperbaiki kekurangan yang ada pada Algoritma K-Means yaitu sensitif terhadap data *outlier* (data dengan nilai-nilai yang ekstrim) [6]. Sebelumnya, banyak peneliti yang telah menggunakan Algoritma K-Medoids ini didalam penelitiannya, diantaranya adalah penelitian Atmaja (2019) yang berjudul “*Implementation of K-medoid clustering algorithm to cluster crime patterns in Yogyakarta*”, penelitian ini membahas tentang proses pengelompokan terhadap pola kejahatan di Yogyakarta [7]. Penelitian Wang dkk (2018) yang berjudul “*K-Medoid Clustering of Data Sequences with Composite Distributions*”, penelitian ini membahas tentang pengelompokan urutan data dengan Distribusi Komposit [8]. Penelitian Yerpude & Dubey (2012) yang berjudul “*Colour Image Segmentation Using K-Medoids Clustering*”, penelitian ini membahas tentang pengelompokan segmentasi citra berwarna [9]. Penelitian Krishnaswamy dkk (2022) yang berjudul “*Application of CRISP-DM methodology for managing human-wildlife conflicts: an empirical case study in India*”, penelitian ini membahas tentang Penerapan CRISP-DM untuk pengelolaan konflik manusia-satwa liar [10]. Penelitian Gunawan (2021) yang berjudul “*Data Mining Using Crisp-Dm Process Framework On Official Statistics: A Case Study Of East Java Province*”, penelitian ini membahas tentang penggunaan CRISP-DM pada data statistik resmi di provinsi Jawa Timur [11].

Algoritma K-Medoids memiliki keunggulan yang signifikan dalam penerapannya dibanding dengan algoritma lain seperti Algoritma K-Means. Sesuai dengan tujuan dibentuknya algoritma K-Medoids, sebagai solusi dalam mengatasi kesensitifan terhadap data *outlier* dan *noise*, Algoritma K-Medoids juga unggul dalam hal kompleksitas komputasi dimana Algoritma K-Medoids ini hanya membutuhkan *pairwise* jarak dari urutan data, yang dapat dihitung sebelumnya [12].

Berdasarkan masalah di atas, dilakukan penelitian dengan menggunakan Algoritma K-Medoids untuk memberikan solusi kepada para pelaku UMKM mengenai bagaimana cara untuk mengatasi krisis ekonomi dan persaingan global bagi para pelaku UMKM dengan menentukan strategi yang baik dan tepat untuk mem-promosikan usahanya. Pada penelitian ini terdapat 3 variabel dan 3 *cluster* yaitu Jenis Usaha, Aset, & Omset yang akan digunakan dan dikelompokkan ke dalam *cluster* tertentu yang memiliki karakteristik yang sama.

2. Metode Penelitian

Pada suatu penelitian, terlebih dahulu dibangun kerangka kerja penelitian. Kerangka kajian ini terdiri dari: menentukan perkara masalah, melakukan studi pustaka, mengumpulkan data, melakukan tingkatan *pre-processing*, menganalisa dan menemukan hasil menggunakan *K-Medoids Clustering*, melaksanakan implementasi menggunakan *Rapid Miner*, dan mengabadikan seluruh proses kegiatan yang dilakukan. Tahap kerja penelitian dapat dilihat seperti pada Gambar 2.



Gambar 2. Tahap Kerja Penelitian

Salah satu metode Data Mining adalah *Clustering*. Makna utama dari *clustering* ini yaitu pengelompokkan sejumlah data atau obyek ke dalam *cluster*/grup sehingga di dalam *cluster* akan berisi data yang sama dengan masing-masing grup [13]. Pada proses *clustering* yang terpenting adalah mengumpulkan pola ke grup yang sesuai untuk menemukan persamaan dan juga perbedaan agar menghasilkan kesimpulan yang berharga.

Clustering ada beberapa *method* yang salah satunya yaitu K-Medoids. Metode K-Medoids ada untuk memperbaiki kelemahan metode K-Means yang sensitive terhadap *outlier* data. Tujuan dari metode K-Medoids yaitu membatasi kedekatan antar titik inti *cluster* & titik data *cluster* [14].

Metode K-Medoids, memiliki beberapa tahap perhitungan yang harus diselesaikan yaitu:

- Memilih nilai *k*, yang dijadikan sebagai jumlah *cluster*
- Tentukan secara random awal medoid dari *n* data
- Menghitung kedekatan per-objek ke medoid sementara menggunakan persamaan Euclidean Distance

$$JED = \sqrt{\sum^n (x_i - y_i)^2} \quad ; 1,2,3, \dots \dots n \quad (1)$$

Dimana *JED* merupakan Jarak Euclidean Distance, *x_i* yaitu nilai medoids pertama pada dimensi satu, *y_i* adalah nilai target cluster pada data cluster pertama dan *n* adalah jumlah target cluster.

- Cari total kedekatan antar cluster dengan menandai jarak minimum obyek ke Medoids.
- Tentukan grup *cluster* pada medoids awal.
- Langkah terakhir yaitu menghitung total simpanan (*S*) dengan cara mencari selisih antara total kedekatan medoids baru dengan total kedekatan medoids lama. Jika total simpanan *S*>0, maka proses *clustering* berhenti, tapi jika *S*<0, maka cari medoids baru lagi hingga didapatkan simpangan *S*>0.

Berikut ini adalah Pseudocode dari K-Medoids Clustering yang disajikan pada Algoritma 1.

Algoritma 1. Pseudocode K-Medoids

KMedoids(*k*, dataset):

```
int Jarak

while true:
    int k medoids as m1, m2, ... mk
    int k cluster as c1, c2, ... ck
    int jarakleast = 0

    for node n in datasets && n not in medoids:
        for m in medoids:
            d = distance(m, n)
            if d = minimum:
                cm.include(n)
                jarakleast = jarakleast + d

    if Jarak > jarakleast:
        return cluster
    else : Jarak = jarakleast
```

3. Hasil dan Pembahasan

Pada tahap berikut ini, dilakukan langkah-langkah dimulai dari menganalisa data sampai dengan proses penyelesaian metode. Adapun 3 hal yang harus dilakukan yaitu sebagai berikut: Analisa Data, Preprocessing Data dan Perhitungan K-Medoids Clustering.

3.1 Analisa Data

Pada penelitian ini, terdapat 3 variabel yang akan digunakan yaitu Jenis Usaha, Asset & Omset. Variabel ini didapatkan dari hasil penyebaran kuisioner kepada para pelaku UMKM Se-Solo Raya dengan total sebanyak 100 data. Ke-3 variabel ini akan diolah dengan menggunakan Algoritma K-Medoids *Clustering*. Sebelum diolah dengan menggunakan Algoritma K-Medoids *Clustering*, terlebih dahulu dilakukan 3 tahapan *Preprocessing* pada Data Mining yaitu *Data Cleaning*, *Transformation Data* & *Normalization Data*. Setelah dilakukan *Preprocessing Data*, mulai dilakukan proses Algoritma K-Medoids dengan menggunakan hasil *Normalization Data* yang sudah di dapatkan sebelumnya pada tahapan *Preprocessing*.

3.2 Preprocessing Data

Preprocessing Data adalah salah satu dari tugas Data Mining yang mencakup persiapan dan transformasi data ke dalam bentuk yang sesuai dengan prosedur Data Mining. *Preprocessing Data* bertujuan untuk memperkecil ukuran data, menemukan relasi antar data, menormalkan data, menghilangkan outlier dan mengekstrak fitur data. *Preprocessing Data* memiliki beberapa teknik dalam pemrosesannya yaitu *Data Cleaning*, *Transformation Data* & *Normalization Data* [15].

3.2.1. Data Cleaning

Data Cleaning berkaitan dengan pendeteksi dan pembersihan data yang *invalid*, salah dan kosong. Pada proses *Data Cleaning*, data akan dicek satu persatu. Bagi data yang memiliki nilai yang *invalid* atau kosong, maka data tersebut akan dihapus [16] [17]. Data yang diolah sebanyak 100 data UMKM Se-Solo Raya. Setelah dilakukan tahap cleaning, didapatkan sebanyak 71 data yang sempurna, lengkap dan tidak rusak. Selanjutnya 71 data UMKM masuk ke langkah berikutnya, yaitu Data Transformation yang disajikan pada Tabel 1.

Tabel 1. Sample Data UMKM

Nama Usaha	Jenis Usaha	Omset	Asset
Ti-A-S	Fashion	Rp.11.000.000 - Rp.15.000.000	≥ Rp.30.000.000
Hekza Cookies	Kuliner	Rp.1.000.000 - Rp.5.000.000	Rp.1.000.000 - Rp.5.000.000
Djiwo Batik Solo	Fashion	Rp.1.000.000 - Rp.5.000.000	Rp.1.000.000 - Rp.5.000.000
Gym2gy	Kerajinan Tangan	Rp.1.000.000 - Rp.5.000.000	Rp.6.000.000 - Rp.10.000.000
Gisela Media	Jasa	Rp.6.000.000 - Rp.10.000.000	≥ Rp.30.000.000
Victorialash	Kecantikan	Rp.1.000.000 - Rp.5.000.000	Rp.6.000.000 - Rp.10.000.000
Dua Putri Gendis	Agribisnis	≤ Rp.1.000.000	≤ Rp.1.000.000

3.2.2. Transformasi Data

Transformation Data juga dikenal sebagai konsolidasi data, dalam proses ini data akan diubah menjadi bentuk yang sesuai untuk prosedur *mining* yaitu berupa data numerik / inisialisasi [18]. Setiap data UMKM ditransformasi ke dalam bentuk numerik.

Tabel 2. Transformasi Data Jenis Usaha UMKM

Jenis Usaha						
Kuliner	Fashion	Jasa	Kecantikan	Kerajinan Tangan	Agribisnis	Lainnya
40	12	8	5	2	2	10
Inisialisasi						
1	2	3	4	5	6	7

Tabel 3. Transformasi Data Asset/Modal UMKM

Modal Awal Usaha	Inisialisasi
≤ Rp.1.000.000	1
Rp.1.000.000 - Rp.5.000.000	2
Rp.6.000.000 - Rp.10.000.000	3
Rp.11.000.000 - Rp.15.000.000	4
Rp.16.000.000 - Rp.20.000.000	5
Rp.21.000.000 - Rp.25.000.000	6
Rp.26.000.000 - Rp.30.000.000	7
≥ Rp.30.000.000	8

Tabel 4. Transformasi Data Omset UMKM

Omset	Inisialisasi
≤ Rp.1.000.000	1
Rp.1.000.000 - Rp.5.000.000	2
Rp.6.000.000 - Rp.10.000.000	3
Rp.11.000.000 - Rp.15.000.000	4
Rp.16.000.000 - Rp.20.000.000	5
Rp.21.000.000 - Rp.25.000.000	6
Rp.26.000.000 - Rp.30.000.000	7
≥ Rp.30.000.000	8

3.2.3. Normalisasi Data

Normalization Data adalah metode peningkatan nilai yang diukur pada skala yang berbeda untuk skala yang sama [19]. Normalisasi data dilakukan dengan cara mengurangi nilai bobot per cluster dengan nilai Min pada cluster tersebut kemudian membaginya dengan hasil pengurangan antara nilai Max dan Min cluster, sehingga didapatkan hasil normalisasi data seperti pada Tabel 5.

Tabel 5. Sampel Normalisasi Data UMKM

Nama Usaha	Jenis Usaha	Omset	Asset
Ti-A-S	0	0,428571429	1
Hekza Cookies	0	0,142857143	0,142857143
Djiwo Batik Solo	0,166666667	0,142857143	0,142857143
Gym2gy	0,666666667	0,142857143	0,285714286
Gisela Media	0,333333333	0,285714286	1
Victorialash	0,5	0,142857143	0,285714286
Dua Putri Gendis	0,833333333	0	0

Selanjutnya, data yang telah di normalisasi dapat digunakan untuk memulai tahapan perhitungan K- Medoids Clustering.

3.3 K-Medoids Clustering

K-Medoid adalah teknik *clustering* yang mempartisi sekumpulan n objek menjadi k jumlah *cluster*. Prinsip dari algoritma ini adalah untuk meminimalkan jumlah ketidaksamaan antara setiap objek dan titik referensi yang sesuai. Dimulai dengan memilih k objek secara acak dari dataset sebagai inisial representasi yang disebut medoid. Rata-rata ketidakmiripan antara medoid dan variabel-variabel dalam minimal *cluster*-nya. Kemudian untuk semua objek lain dalam dataset, ia menetapkan ke *cluster* terdekat tergantung pada jarak benda dan medoid. Setelah itu, medoid baru dipilih [20].

3.3.1. Memilih nilai k, jumlah cluster

Data UMKM dibagi menjadi 3 cluster, yaitu: cluster 1 jenis usaha, cluster 2 omset, dan cluster 3 asset.

3.3.2. Tentukan secara random awal medoid dari n data

Pilih medoid pertama dari semua cluster secara random dari data normalisasi yang disajikan pada Tabel 6.

Tabel 6. Medoid Pertama

Bidang Usaha	Pendapatan	Modal Awal Usaha
0	0,142857143	1
1	0,714285714	1
0	0,142857143	0,285714286

3.3.3. Menghitung kedekatan per-objek ke medoid sementara

Untuk menghitung kedekatan per-objek, digunakan persamaan Euclidean Distance pada Persamaan 1. Berikut ini adalah contoh perhitungan pada data pertama.

a. Data pertama ke pusat cluster 1.

$$JED(X_1C_1) = \sqrt{(0-0)^2 + (0,42-0,14)^2 + (1-1)^2} = 0,285$$

b. Data pertama ke pusat cluster 2.

$$JED(X_1C_2) = \sqrt{(0-1)^2 + (0,42-0,71)^2 + (1-1)^2} = 1,040$$

c. Data pertama ke pusat cluster 3.

$$JED(X_1C_3) = \sqrt{(0-0)^2 + (0,42-0,14)^2 + (1-0,28)^2} = 0,769$$

Lakukan proses serupa untuk data ke-2 sampai data terakhir, sehingga akan dihasilkan jarak masing-masing objek ke medoids sementara seperti pada Tabel 7.

Tabel 7. Jarak Objek Ke Medoid Awal

C1	C2	C3	Jarak	Cluster
0,2857142	1,0400157	0,7693092	0,2857142	1
0,7346938	1,4356965	1	0,7346938	1
0,9013605	1,3250165	1,0137937	0,9013605	1
1,1768707	0,9735737	1,2018504	0,9735737	2
0,3626558	0,7925389	1,0637289	0,3626558	1
1,8454895	1,2401539	1,3095238	1,2401539	2
...	...	...	...	...
1,0102040	1,0424656	1,1180339	1,0102040	1
Total Kedekatan Jarak Min			56,411267	

3.3.4. Cari total kedekatan antar cluster dengan menandai jarak minimum obyek ke Medoids.

Berdasarkan jarak minimum objek ke medoids awal / sementara pada tabel 6, maka dapat dihitung total kedekatannya dengan cara menjumlahkan keseluruhan jarak minimal di tabel 5, sehingga diperoleh nilai kedekatannya sebanyak 56,411267.

3.3.5. Lakukan iterasi medoids baru dengan memilih kembali secara acak medoids baru yang disajikan pada Tabel 8.

Tabel 8. Medoid Ke-2

Bidang Usaha	Pendapatan	Modal Awal Usaha
0	0,428571429	1
0	0,142857143	0,142857143
0	0,142857143	0,142857143

Ulangi kembali langkah pada 3.3.3. sampai 3.3.4., sehingga didapatkan total kedekatannya sebesar 23,7847243.

3.3.6. Cari nilai total simpangan (S).

Selanjutnya, menghitung total simpangan (S) dengan cara menghitung selisih antar jumlah kedekatan medoid baru dengan jumlah kedekatan medoid lama. Apabila $S > 0$, maka proses dihentikan. Jika $S < 0$, cari medoid baru hingga menemukan total $S > 0$.

$$S = 23,7847243 - 56,411267 = -32,626543$$

Dari nilai S di atas, didapat hasil negatif yang ditandai dengan nilai $S < 0$. Oleh sebab itu, tentukan medoids baru secara random dan lakukan proses dan langkah 3.3.3. sampai 3.3.4.

Tabel 9. Medoid Ke-3

Bidang Usaha	Pendapatan	Modal Awal Usaha
0,666666667	0,142857143	0,285714286
0	0,142857143	0,142857143
0	0	0,142857143

Ulangi kembali langkah pada 3.3.3. sampai 3.3.4., sehingga didapatkan total kedekatannya sebesar 20,0388488.

Selanjutnya, hitung kembali total simpangan (S).

$$S = 20,0388488 - 23,7847243 = -3,7458755$$

Total simpangan pada iterasi ke-3 masih bernilai – (<0), oleh karena itu dilakukan iterasi medoid ke-4 dengan medoid seperti pada tabel 10.

Tabel 10. Medoid Ke-4

Bidang Usaha	Pendapatan	Modal Awal Usaha
0,333333333	0,285714286	0,142857143
0	0,142857143	0,285714286
1	0,714285714	1

Ulangi kembali langkah pada 3.3.3. sampai 3.3.4., sehingga didapatkan total kedekatannya sebesar 21,3004979.

Selanjutnya, hitung kembali total simpangan (S).

$$S = 21,3004979 - 20,0388488 = 1,26164909$$

Dari nilai S di atas, didapat bahwa $S > 0$. Dikarenakan nilai $S > 0$, maka tahapan clustering dihentikan.

Tabel 11. Data Sampel Hasil Cluster

Nama Usaha	Bidang Usaha	Pendapatan	Modal Awal Usaha	Cluster
Ti-A-S	1	4	8	2
Hekza Cookies	1	2	2	2
Djiwo Batik Solo	2	2	2	1
Gym2gy	5	2	3	1
Gisela Media	3	3	8	3
Victorialash	4	2	3	1
...	...	...	...	...
Dua Putri Gendis	6	1	1	1

Berdasarkan hasil perhitungan metode K-Medoids, diketahui hasil pada cluster 1 sebesar 25 data, cluster 2 sebesar 39 data & cluster 3 sebesar 7 data yang disajikan pada tabel 12, 13 & 14.

Tabel 12. Data Sampel Hasil Cluster 1

Cluster 1
Cluster 1 : 25 data dari berbagai jenis usaha
Kuliner : 3
Fashion : 9
Jasa : 5
Kecantikan : 3
Kerajinan Tangan : 1
Agribisnis : 1
Lainnya : 3
Dengan rata-rata omset dan asset sama yaitu antara Rp.1.000.000 - Rp.5.000.000

Tabel 13. Data Sampel Hasil Cluster 2

Cluster 2
Cluster 2 : 39 data dari berbagai jenis usaha
Kuliner : 37
Fashion : 2
Jasa : -
Kecantikan : -
Kerajinan Tangan : -
Agribisnis : -
Lainnya : -
Dengan rata-rata omset dan asset sama yaitu antara Rp.1.000.000 - Rp.5.000.000

Tabel 14. Data Sampel Hasil Cluster 3

Cluster 3
Cluster 2 : 7 data dari berbagai jenis usaha
Kuliner : -
Fashion : -
Jasa : 1
Kecantikan : -
Kerajinan Tangan :-
Agribisnis : -
Lainnya : 6
Dengan rata-rata omset antara Rp.21.000.000 - Rp.25.000.000 dan asset sebesar $\geq$ Rp.30.000.000

Dari tabel 12, 13 & 14 di atas, diketahui bahwa cluster 1 memiliki rata-rata asset dan omset antara Rp.1.000.000 - Rp.5.000.000 yang didominasi oleh jenis usaha pada bidang fashion. Cluster 2 memiliki rata-rata asset sebesar $\geq$ Rp.30.000.000 dan omset antara Rp.1.000.000 - Rp.5.000.000 yang didominasi oleh jenis usaha pada bidang kuliner. Cluster 3 memiliki rata-rata asset sebesar $\geq$ Rp.30.000.000 dan omset antara Rp.21.000.000 - Rp.25.000.000 yang didominasi oleh jenis usaha pada bidang lainnya (diluar dari ke-6 bidang usaha lain).

Dari hasil analisa cluster data UMKM di atas, dapat dilihat bahwa cluster 2 merupakan data yang paling banyak muncul dengan 39 data UMKM. Data pada cluster 2 & cluster 1 merupakan data centroid yang dipilih berdasarkan omset usaha yang paling kecil. Meskipun begitu, jika dilihat dari jumlah UMKM terbanyak terdapat pada cluster 2, maka berdasarkan hal tersebut cluster ke-2 dipilih sebagai data centroid dengan omset yang paling kecil. Selanjutnya, hasil analisa cluster dapat dihubungkan dengan promotion strategy yang dikutip dari penelitian Astuti dkk pada tahun 2019 seperti pada tabel 15.

Tabel 15. Promotion Strategy (Astuti dkk, 2019)

No	Strategy	C1	C2	C3
1	Mengikuti kegiatan temu usaha UMKM	√	√	√
2	Mengikuti pameran UMKM	√	√	
3	Bergabung dengan pojok UMKM		√	
4	Memanfaatkan dunia digital sebagai upaya pengembangan usaha	√	√	√

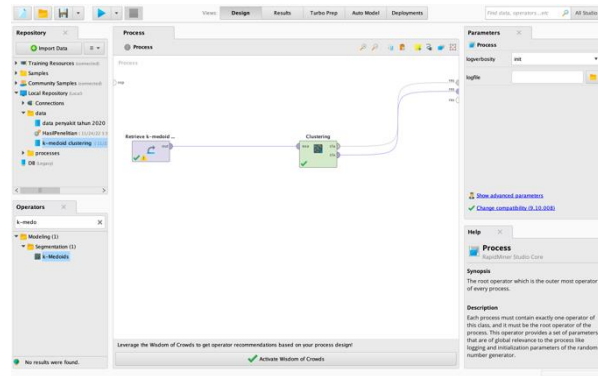
Tabel 15 menunjukkan bahwa cluster 2 disarankan untuk mengikuti ke-4 promotion strategy yang ada. Hal ini dikarenakan cluster 2 memiliki jumlah UMKM yang paling tinggi yaitu 39 UMKM dengan rata-rata omset dan asset paling rendah. Untuk cluster 1, rata-rata omset dan asset sama dengan cluster ke-2, akan tetapi jumlah UMKM yang tergolong pada cluster 1 lebih sedikit sehingga untuk cluster 1 cukup mengikuti 3 promotion strategy diantaranya: mengikuti kegiatan temu usaha UMKM, mengikuti pameran UMKM & memanfaatkan dunia digital sebagai upaya pengembangan usaha. Cluster ke-3 memiliki rata-rata omset dan asset yang paling tinggi diantara cluster 1 dan 2 sehingga cluster ke-3 cukup mengikuti 2 promotion strategy diantaranya: mengikuti kegiatan temu usaha UMKM & memanfaatkan dunia digital sebagai upaya pengembangan usaha. Selanjutnya dilakukan proses analisa akhir promotion strategy berdasarkan promotion mix yang dapat dilihat pada tabel 16.

Tabel 16. Promotion Strategy Berdasarkan Promotion Mix

No	Promotion Strategy Berdasarkan Promotion Mix	C1	C2	C3
1	Periklanan			√
2	Penjualan Personal	√	√	√
3	Promosi Penjualan	√	√	√
4	Hubungan Masyarakat	√	√	√
5	Pemasaran Langsung	√	√	√

Berdasarkan hasil analisa akhir cluster dan promotion mix, diketahui bahwa cluster 1 & 2 dengan rata-rata asset dan omset paling kecil memungkinkan untuk UMKM tersebut melakukan penjualan personal, promosi penjualan, hubungan masyarakat dan pemasaran langsung. Untuk cluster 3 dengan rata-rata asset dan omset paling tinggi memungkinkan untuk melakukan keseluruhan aktivitas promotion strategy.

3.3.7 Penerapan Aplikasi RapidMiner Studio



Gambar 3. Penghubung Data dan K- Medoids Clustering

The screenshot displays the Orange Data Mining software interface. The top menu bar includes File, View, Design, Results, Turbo Fan, Auto Model, and Deployments. Below the menu is a toolbar with icons for opening files, saving, undo, redo, and running models. The main window is titled "Cluster Model (Clustering)" and contains a table of results.

No	ID	cluster	Belong Uns.	Prediction	Model Acc.
1	cluster_1	0	0.143	0	
2	cluster_1	0	0.143	0.143	
3	cluster_1	0	0.143	0.143	
4	cluster_2	1	0.143	0.143	
5	cluster_0	0.143	0.143	0.143	
6	cluster_1	0	0.143	0.429	
7	cluster_1	0	0	0	
8	cluster_1	0	0.143	0.143	
9	cluster_0	0.047	0.143	0.286	
10	cluster_1	0	0.143	0.143	
11	cluster_1	0	0	0.143	
12	cluster_2	0	0	0.143	
13	cluster_1	0	0.143	0.143	
14	cluster_1	0	0.143	0.143	
15	cluster_1	0	0.286	0.286	
16	cluster_2	0.333	0.286	1	
17	cluster_1	0	0.143	0.143	
18	cluster_2	0	0.429	0.286	
19	cluster_0	0.333	0.143	0	
20	cluster_1	0	0.143	0.143	
21	cluster_2	0	0.143	0.143	
22	cluster_2	1	1	0.286	

Example(1) 7 examples, 2 spatial attributes, 1 regular attribute

On the right side, there is a sidebar with sections for "Repository" (listing Training Resources like Sample, Community Samples, Local Repository, Connections, and Data) and "Data" (listing recent datasets like data_perspek_tahun_2020, northwestern, and model_clustering).

Gambar 4. Nilai Cluster Data RapidMiner

The screenshot displays the Orange3 data mining environment. The main workspace contains a workflow with three data files (data1.csv, data2.csv, data3.csv) connected to a 'Cluster' widget. The 'Cluster' widget's output is connected to a 'Table' widget. The 'Table' widget displays the results of the clustering, showing columns for 'Cluster', 'X', 'Y', and 'Z'. The 'Cluster' column shows three distinct clusters of data points. The 'Table' widget is set to 'Table' view, showing a list of data points with their assigned cluster numbers. The 'Cluster' widget's settings are visible on the right, showing 'K-Means' as the selected algorithm and '3' as the number of clusters. The 'Table' widget's settings are also visible, showing 'Table' as the selected view and 'All' as the selected columns.

Gambar 5. Nilai Cluster Text RapidMiner

4. Kesimpulan

Hasil Analisa Promotion Strategy UMKM dilakukan dengan membentuk clustering dari data UMKM Se-Solo Raya menggunakan 3 cluster utama sebagai objek penelitian yaitu: jenis usaha, omset & aset. Data ini diolah menggunakan algoritma k-medoids clustering dan diimplementasikan menggunakan rapid miner. Hasil clustering pada algoritma ini akan diselaraskan dan disesuaikan dengan kategori promotion mix. Hasil penyelarasan kedua metode ini berupa strategi promosi seperti apa yang tepat untuk dilakukan oleh pelaku UMKM yang didasari dari jenis usaha, omset dan aset.

Daftar Rujukan

- [1] Aprianto, E. (2019). Empowering the Amangtiwi-UMKM in Malang through Basic English Language Skill. *Journal Community Development and Society*, 1(1). <https://doi.org/10.25139/cds.v1i1.1625>
- [2] Astuti, D. (2019). Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(2), 60–72. <https://doi.org/10.20895/inista.v1i2.71>

- [3] Rak, T., & Żyła, R. (2022). Using Data Mining Techniques for Detecting Dependencies in the Outcoming Data of a Web-Based System. *Applied Sciences*, 12(12), 6115. <https://doi.org/10.3390/app12126115>
- [4] Ashish B., Sumit G., Shreyas K & Suhasini P. (2018). Study of Data Mining Concept. *International Journal of New Innovations in Engineering and Technology*. Volume 9 Issue 1. ISSN : 2319-6319.
- [5] Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-Means Clustering Algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/access.2020.2988796>
- [6] Dinata, R. K., Retno, S., & Hasdyna, N. (2021). Minimization of the Number of Iterations in K-Medoids Clustering with Purity Algorithm. *Revue d'Intelligence Artificielle*, 35(3), 193–199. <https://doi.org/10.18280/ria.350302>
- [7] Atmaja, E. H. S. (2019). Implementation of k-Medoids Clustering Algorithm to Cluster Crime Patterns in Yogyakarta. *International Journal of Applied Sciences and Smart Technologies*, 1(1), 33–44. <https://doi.org/10.24071/ijasst.v1i1.1859>
- [8] Wang, T., Li, Q., Bucci, D. J., Liang, Y., Chen, B., & Varshney, P. K. (2019). K-Medoids Clustering of Data Sequences With Composite Distributions. *IEEE Transactions on Signal Processing*, 67(8), 2093–2106. <https://doi.org/10.1109/tsp.2019.2901370>
- [9] Yerpude, A., & Dubey, S. (2012). Colour image segmentation using K-medoids clustering. *Int J Comput Technol Appl*, 3(1), 152–4.
- [10] Krishnaswamy, V., Singh, N., Sharma, M., Verma, N., & Verma, A. (2022). Application of CRISP-DM methodology for managing human-wildlife conflicts: an empirical case study in India. *Journal of Environmental Planning and Management*, 1–27. <https://doi.org/10.1080/09640568.2022.2070460>
- [11] Gunawan, G. (2021). Data Mining Using Crisp-Dm Process Framework On Official Statistics: A Case Study Of East Java Province. *Jurnal Ekonomi Dan Pembangunan*, 29(2), 183–198. <https://doi.org/10.14203/Jep.29.2.2021.183-198>
- [12] Wang, T., Li, Q., Bucci, D. J., Liang, Y., Chen, B., & Varshney, P. K. (2019). K-Medoids Clustering of Data Sequences With Composite Distributions. *IEEE Transactions on Signal Processing*, 67(8), 2093–2106. <https://doi.org/10.1109/tsp.2019.2901370>
- [13] Yuda Irawan. (2019). Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (Jtiulm)*, 4(1), 13–20. <https://doi.org/10.20527/Jtiulm.V4i1.34>
- [14] Rahman, F., Ridho, I. I., Muflih, M., Pratama, S., Raharjo, M. R., & Windarto, A. P. (2020). Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country. *IOP Conference Series: Materials Science and Engineering*, 835(1), 012058. <https://doi.org/10.1088/1757-899x/835/1/012058>
- [15] Alasadi, S. A., & Bhaya, W. S. (2017). Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences*, 12(16), 4102–4107.
- [16] Rahm, E., & Do, H. H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, 23(4), 3–13.
- [17] Patil, R. Y., & Kulkarni, R. V. (2012). A review of data cleaning algorithms for data warehouse systems. *International Journal of Computer Science and Information Technologies*, 3(5), 5212–5214.
- [18] Li, Z., & Liu, L. (2022). Project Cost Management Information Solution Based on Data Mining Technology. 2022 3rd Asia-Pacific Conference on Image Processing, Electronics and Computers. <https://doi.org/10.1145/3544109.3544363>
- [19] Sinsomboonthong, S. (2021). Efficiency Comparison in Prediction of Normalization with Data Mining Classification. *Advances in Science, Technology and Engineering Systems Journal*, 6(4), 130–137. <https://doi.org/10.25046/aj060415>
- [20] Hudaib, A., Khanafseh, M., & Surakhi, O. (2018). An Improved Version of K-medoid Algorithm using CRO. *Modern Applied Science*, 12(2), 116. <https://doi.org/10.5539/mas.v12n2p116>