

Implementasi Metode K-Means Clustering dalam Analisis Persebaran UMKM di Jawa Barat

Nursyifa¹, Resti Noor Fahmi²

Fakultas Ilmu Komputer, Program Studi Teknik Informatika
Universitas Singaperbangsa Karawang

e-mail: ¹nursyifa17168@student.unsika.ac.id, ²resti.noor17019@student.unsika.ac.id

Diterima: 09 Oktober 2021; Direvisi: 04 Nopember 2021; Disetujui: 23 Nopember 2021

Abstrak

Usaha Mikro Kecil dan Menengah atau UMKM ialah usaha produktif yang sudah teruji membuka lapangan kerja dan menjadi penggerak roda perekonomian di Indonesia. Perlu adanya pengembangan potensi dalam melaksanakan UMKM dengan menganalisis strategi persebaran UMKM dan peningkatan jumlah UMKM. Penelitian ini akan menjelaskan cara mengimplementasikan metode k-means clustering untuk menganalisis persebaran UMKM sehingga diharapkan dapat menjadi perhatian bagi pemerintah atau institusi terkait dalam meningkatkan perekonomian UMKM di Jawa Barat. Clustering UMKM akan dibagi menjadi 3 bagian berdasarkan tingkat persebaran UMKM yaitu cluster 0 atau rendah memperoleh hasil 9 Kabupaten/Kota, cluster 1 atau sedang memperoleh hasil 15 Kabupaten/Kota, dan cluster 2 atau tinggi memperoleh hasil 3 Kabupaten/Kota. Hasil evaluasi clustering akan dibandingkan dengan 2 metode evaluasi yaitu silhouette coefficient yang menghasilkan nilai sebesar 0,73, sedangkan metode davies bouldin index (DBI) menghasilkan nilai sebesar 0,29. Dari hasil perbandingan kedua evaluasi cluster tersebut menunjukkan hasil cluster dengan menerapkan algoritma k-means yang terbentuk adalah klaster yang terbaik.

Kata kunci: davies bouldin index, jawa barat, k-means clustering, silhouette coefficient, umkm

Abstract

Micro Small and Medium Enterprises or MSME are productive businesses that have been tested to opening jobs and become the cog in the economy in Indonesia. There needs to be potential development in MSMEs by analyzing MSME distribution strategies and an increase in the number of MSMEs. This research will explain how to implement the k-means clustering method to analyze the distribution of MSMEs so that it is expected to be a concern for the government or related institutions in improving the MSME economy in West Java. Clustering MSMEs will be divided into 3 parts based on the distribution level of MSMEs, namely cluster 0 or low to obtain results of 9 Regencies/Cities, cluster 1 or are obtaining results of 15 Regencies/Cities, and clusters 2 or high obtain results of 3 Regencies/Cities.

Keywords: davies bouldin index, k-means clustering, msme, silhouette coefficient, west java

1. PENDAHULUAN

Kesejahteraan masyarakat suatu negara salah satunya bergantung kepada kestabilan perekonomiannya. Ada beberapa hal suatu negara dikatakan stabil dalam ekonominya, diantaranya nilai mata uang rupiah stabil, laju pertumbuhan ekonomi yang naik, meratanya pendapatan masyarakat, tidak terjadi inflasi maupun deflasi dan lain-lain. Contohnya di Jawa

Barat, laju ekonominya berada di kuartal I tahun 2021. Dalam hal ini stabilitas keuangan Jawa Barat ini dapat dikatakan masih dalam kondisi baik. Namun, dilihat dari Tingkat Pengangguran Terbuka (TPT) di Jawa Barat ini terbilang cukup tinggi yaitu sebanyak 8,92%. Hal ini berkaitan dengan kesejahteraan masyarakat yang belum pulih seutuhnya. Pada september 2020, tingkat kemiskinan mengalami peningkatan dibandingkan tahun 2019 menjadi 8,43%. Namun, angka tersebut belum memperlihatkan adanya perbaikan kondisi ekonomi yang terjadi di Jawa Barat sejak tahun 2020 sampai triwulan I 2021 [1]. Perekonomian Jawa Barat diprediksi di tahun 2021 ini akan kembali tumbuh positif sesudah berkontraksi dalam akibat dari pandemi covid-19. Untuk dapat mewujudkan prediksi dan meningkatkan laju ekonomi di Jawa Barat salah satunya yaitu dengan melakukan upaya pemerataan pendapatan masyarakat. Upaya ini dapat dilakukan dengan meningkatkan jumlah dan persebaran UMKM di Jawa Barat.

Usaha Mikro Kecil dan Menengah atau UMKM ialah usaha produktif yang sudah teruji membuka lapangan pekerjaan dan menjadi penggerak roda perekonomian di Indonesia [2]. UMKM merupakan aktivitas usaha yang dikerjakan oleh individu atau kelompok yang mempunyai tujuan menyejahterakan individu ataupun kelompoknya. UMKM memiliki fungsi strategis dalam pengembangan ekonomi nasional [3]. UMKM berperan penting dalam salah satu faktor pendukung perekonomian Indonesia.

Untuk dapat mendukung mewujudkan struktur perekonomian nasional yang seimbang, membangun kemandirian ekonomi masyarakat, penyerapan tenaga kerja yang maksimal, mengurangi tingkat kemiskinan, pertumbuhan ekonomi dan pemerataan pendapatan masyarakat, sehingga diperlukan perhatian terhadap perkembangan UMKM di Jawa Barat ini. Betapa pentingnya peran UMKM dalam membangun negeri dan juga mewujudkan ekonomi yang berkesinambungan di Jawa Barat, maka penelitian ini akan menganalisis persebaran UMKM di Jawa Barat sehingga diharapkan dapat menjadi perhatian bagi pemerintah atau institusi terkait dalam memberikan solusi dan meningkatkan jumlah serta persebaran UMKM di Jawa Barat.

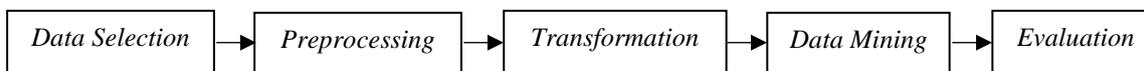
Penelitian sebelumnya yang dilakukan oleh Astuti dkk ialah mengelompokkan data yang sesuai atau klusterisasi sebagai dasar penentu strategi promosi dengan metode CRISP-DM dan algoritma *K-Means Clustering* yang menghasilkan 3 *cluster* [5]. Aristika & Hartono menerapkan *Clustering K-Means* dalam penentuan pengaruh media sosial facebook terhadap UMKM di Kecamatan Pekanbaru Kota yang memperoleh dua kelompok pemakai facebook yaitu berpengaruh dan sangat berpengaruh [6]. Penelitian Fitriani & Elisa memanfaatkan *data mining* untuk menentukan media promosi UMKM di Kota Batam menggunakan metode KDD dan algoritma *Clustering K-means* yang diuji dengan *software* RapidMiner 5.3 menghasilkan 3 *cluster* yaitu kurang diminati, paling tidak diminati dan paling diminati [7]. Setiawan melakukan penelitian dengan menerapkan *data mining* dalam menentukan strategi promosi mahasiswa baru menggunakan metode CRISP-DM dan algoritma *K-Means Clustering* untuk menghasilkan profil yang mempunyai kesamaan atribut sehingga memperoleh jumlah kluster 4 ($k=4$) dengan bantuan aplikasi RapidMiner 7.0 [8].

Maka penelitian ini untuk menganalisis persebaran UMKM di Jawa Barat akan mengimplementasikan metode *K-Means Clustering* untuk proses pengelompokan daerah yang memiliki persebaran atau jumlah UMKM yang rendah, sedang hingga tinggi. Pada implementasi metode Analisis *K-Means Cluster*, data yang dapat diolah pada perhitungan merupakan data angka atau numerik. Masing-masing data dihitung dari kedekatan dengan nilai *centroid* yang sebelumnya telah ditentukan, jarak terkecil antara data dengan tiap-tiap *centroid* ialah anggota *cluster* yang terdekat [4]. Perbedaan penelitian ini dengan sebelumnya adalah pada evaluasi kualitas penerapan *clustering* yang akan dilakukan menggunakan perbandingan dua evaluasi yaitu dengan metode *silhouette coefficient* dan metode *davies bouldin index* (DBI) untuk mengetahui perbandingan hasil kualitas *cluster* dari pengimplementasian algoritma *k-means*.

2. METODE PENELITIAN

Penelitian ini menerapkan metodologi *Knowledge Discovery in Database* (KDD). KDD

merupakan ekstraksi informasi potensial, dan implisit dalam proses non-trivial dalam mencari dan mendefinisikan pola (*pattern*) dalam data [9]. Tahapan metodologi yang akan dilakukan pada penelitian ini, digambarkan dalam sebuah alur metodologi penelitian seperti pada gambar 1.



Gambar 1. Alur metodologi penelitian

2.1. Data Selection

Pada tahap ini dilakukan proses pemilihan, pencarian data dan pengambilan data yang akan digunakan dari sekumpulan data operasional. Pengambilan data diambil dari data *public* melalui situs opendata.jabarprov.go.id diperoleh Dinas Perindustrian dan Perdagangan yang dipublikasikan tiap periode setahun sekali. Data yang digunakan adalah data jumlah unit industri kecil, menengah, dan besar yang ada di Jawa Barat berdasarkan Kabupaten/Kota dari tahun 2015 sampai dengan 2019.

2.2. Preprocessing

Preprocessing adalah proses pembersihan dan penyederhanaan teks [10]. Dalam proses pengambilan data, data yang didapatkan pasti berupa *unstructured* data, yang di mana isi dalam data tersebut masih terdapat *noise*. Maka dari itu, tujuan dari *preprocessing* ini untuk menghilangkan *noise* yang ada. Selain itu, proses pembersihan data lainnya diantaranya menghilangkan duplikasi data, mengoreksi kesalahan pada data, dan meninjau data yang inkonsisten [11].

2.3. Transformation

Tahap transformasi adalah proses pengembangan *dataset* sesuai dengan kebutuhan pada tahap *modeling* atau tahap *data mining*, karena sebagian metode dalam *data mining* membutuhkan ketentuan khusus sebelum diolah dalam *data mining* [12].

2.4. Data Mining

Tahap *data mining* merupakan proses di mana pengolahan data, proses mencari pola atau informasi berdasarkan metode ataupun teknik yang akan digunakan. Pada pemilihan metode yang sesuai digunakan sangat bergantung kepada tujuan dan proses KDD keseluruhan [11].

2.5. Evaluation

Tahap evaluasi yaitu untuk mengetahui sejauh mana kualitas dari penelitian yang dilakukan yang merupakan interpretasi hasil pemodelan yang digunakan [13]. Pada penelitian ini dilakukan perbandingan evaluasi hasil *clustering* yaitu dengan menggunakan metode *Silhouette Coefficient* dan *Davies Bouldin Index*.

Metode DBI adalah teknik evaluasi pada *clustering* dengan fungsi rasio dari jumlah antara *cluster scatter* hingga *cluster separation*. DBI didasarkan pada kemiripan dari ukuran *cluster* berdasar pada penyebaran *cluster* dan ketidaksamaan ukuran *cluster* [14]. Prinsip pendekatan pada pengukuran evaluasi DBI ialah memaksimalkan jarak *inter cluster* dan meminimalkan jarak *intra cluster* [15]. Semakin kecil nilai DBI yang dihasilkan maka skema klaster termasuk ke dalam *cluster* yang optimal [16].

Metode *silhouette coefficient* yaitu salah satu teknik evaluasi untuk mengetahui kekuatan dan kualitas cluster serta seberapa baik suatu objek ditempatkan pada suatu cluster [17].

3. HASIL DAN PEMBAHASAN

Implementasi *clustering k-means* dilakukan pada data UMKM yang terdapat di Jawa Barat terdiri dari 27 Kabupaten/Kota.

3.1. Data Selection

Pada tahap *data selection* dilakukan proses pencarian data dan pengambilan data yang akan digunakan. Pengambilan data diambil dari data *public* melalui situs opendata.jabarprov.go.id diperoleh Dinas Perindustrian dan Perdagangan yang dipublikasikan tiap periode setahun sekali. Data yang digunakan adalah data jumlah unit industri kecil, menengah, dan besar yang ada di Jawa Barat berdasarkan Kabupaten/Kota dari tahun 2015 sampai dengan 2019.

3.2. Preprocessing

Preprocessing adalah proses pembersihan dan perubahan data mentah menjadi data yang siap digunakan pada tahap berikutnya. *Dataset* yang didapatkan berupa data yang terdiri dari 135 *record*. Pada data di bawah terdapat data id, kode provinsi, nama provinsi, kode kabupaten kota, nama kabupaten kota, jumlah unit, satuan serta tahun. Kode provinsi untuk wilayah Provinsi Jawa Barat ialah 32. Berikut detail isi dataset awal yang terdapat pada tabel 1.

Tabel 1. *Dataset* awal

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	jumlah_unit	satuan	tahun
0	32	Jawa Barat	3201	Kabupaten Bogor	15242	Unit	2015
1	32	Jawa Barat	3202	Kabupaten Sukabumi	15801	Unit	2015
2	32	Jawa Barat	3203	Kabupaten Cianjur	1257	Unit	2015
3	32	Jawa Barat	3204	Kabupaten Bandung	13999	Unit	2015
4	32	Jawa Barat	3205	Kabupaten Garut	9846	Unit	2015
.....
131	32	Jawa Barat	3276	Kota Depok	10549	Unit	2019
132	32	Jawa Barat	3277	Kota Cimahi	6181	Unit	2019
133	32	Jawa Barat	3278	Kota Tasikmalaya	10005	Unit	2019
134	32	Jawa Barat	3279	Kota Banjar	9821	Unit	2019

Tabel 1 di atas memperlihatkan *dataset* jumlah UMKM yang ada di Jawa Barat dengan terdapat beberapa atribut data. Pada penelitian ini atribut yang digunakan hanya nama kabupaten kota, jumlah unit, dan tahun. Maka dari itu pada proses *preprocessing* akan dilakukan penghapusan atribut yang tidak dibutuhkan, dapat dilihat pada gambar 2 berikut.

```
df2=df.drop(["kode_provinsi","nama_provinsi","kode_kabupaten_kota","satuan"], axis=1)
```

Gambar 2. Proses menghapus atribut yang tidak dibutuhkan

Pada tabel 2 di bawah merupakan hasil dari penghapusan atribut yang tidak dibutuhkan sehingga hanya terdapat atribut id, nama_kabupaten_kota, jumlah_unit, dan tahun. Di mana data tersebut digunakan pada proses selanjutnya.

Tabel 2. Data setelah proses penghapusan atribut yang tidak dibutuhkan

id	nama_kabupaten_kota	jumlah_unit	tahun
0	Kabupaten Bogor	15242	2015
1	Kabupaten Sukabumi	15801	2015

id	nama_kabupaten_kota	jumlah_unit	tahun
2	Kabupaten Cianjur	1257	2015
3	Kabupaten Bandung	13999	2015
4	Kabupaten Garut	9846	2015
.....
131	Kota Depok	10549	2019
132	Kota Cimahi	6181	2019
133	Kota Tasikmalaya	10005	2019
134	Kota Banjar	9821	2019

3.3. Transformation

Tahap ini yaitu melakukan perubahan data untuk memudahkan pada proses pengolahan data sesuai dengan kebutuhan pada proses *data mining*. Perubahan data dilakukan dengan mengubah yang tadinya data dalam satu kolom menunjukkan jumlah UMKM tiap tahunnya, pada tahap ini mengubah jumlah UMKM tersebut menjadi beberapa kolom yang dikelompokkan berdasarkan tahun. Kemudian menambahkan atribut jumlah yang menunjukkan jumlah UMKM selama tahun 2015 s.d. 2019.

Transformasi data ini menghasilkan 27 data yang berisi nama kabupaten/kota di Jawa Barat yang terdapat pada kolom Kabupaten/Kota. Pada setiap kabupaten/kota terdapat data jumlah UMKM sesuai tahunnya dari tahun 2015 hingga 2019 yang terletak pada kolom 2015, 2016, 2017, 2018, dan 2019. Pada kolom terakhir yaitu jumlah merupakan nilai keseluruhan dari jumlah UMKM di Jawa Barat. Pada tabel 3 berikut menunjukkan data jumlah UMKM di Jawa Barat hasil dari proses transformasi data.

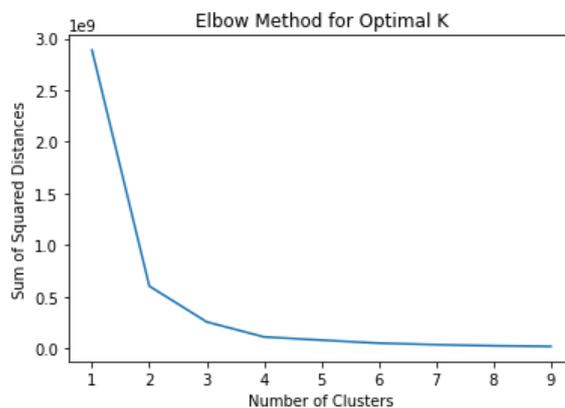
Tabel 3. Data jumlah UMKM di Jawa Barat

No.	Kabupaten/Kota	2015	2016	2017	2018	2019	Jumlah
1	Kabupaten Bogor	15242	15711	14204	15961	16049	77167
2	Kabupaten Sukabumi	15801	15955	12556	16165	16166	76643
3	Kabupaten Cianjur	1257	1300	1526	1335	1346	6764
4	Kabupaten Bandung	13999	14094	15917	14229	14233	72472
5	Kabupaten Garut	9846	9880	2506	9993	10000	42225
6	Kabupaten Tasikmalaya	1483	1500	495	1526	1551	6555
7	Kabupaten Ciamis	1423	1476	9981	1534	1539	15953
8	Kabupaten Kuningan	2461	2481	5232	2534	2573	15281
9	Kabupaten Cirebon	10853	10898	1503	10942	11144	45340
10	Kabupaten Majalengka	7428	7446	2427	7473	7475	32249
11	Kabupaten Sumedang	5151	5204	352	5254	5254	21215
12	Kabupaten Indramayu	2392	2408	10938	2429	2430	20597
13	Kabupaten Subang	3433	3454	10515	3498	3538	24438
14	Kabupaten Purwakarta	10962	11019	11098	11125	11125	55329
15	Kabupaten Karawang	10113	10332	7457	10548	10553	49003
16	Kabupaten Bekasi	11153	11704	1324	12639	12653	49473
17	Kabupaten Bandung Barat	214	238	15989	384	389	17214
18	Kabupaten Pangandaran	486	487	3487	495	531	5486
19	Kota Bogor	8494	8829	9446	8875	8875	44519
20	Kota Sukabumi	9462	9530	9987	9571	9571	48121

No.	Kabupaten/Kota	2015	2016	2017	2018	2019	Jumlah
21	Kota Bandung	10901	10943	8867	11124	11124	52959
22	Kota Cirebon	9421	9428	10538	9450	9452	48289
23	Kota Bekasi	10002	10142	11092	10418	10433	52087
24	Kota Depok	10436	10483	6176	10540	10549	48184
25	Kota Cimahi	6155	6158	10360	6181	6181	35035
26	Kota Tasikmalaya	9801	9938	9767	10005	10005	49516
27	Kota Banjar	9712	9719	9551	9798	9821	48601

3.4. Data Mining

Tahap ini yaitu proses pengolahan data sehingga didapatkan informasi atau pengetahuan menggunakan metode tertentu. Metode yang digunakan ialah *clustering* dengan menggunakan algoritma *k-means*. Pengolahan data diolah menggunakan bahasa pemrograman *python*. Di mana tahapan awal pada proses *clustering k-means* yaitu menentukan jumlah *cluster* yang akan dibentuk. Penentuan jumlah *cluster* yang ini dilakukan dengan metode *elbow* guna menentukan optimasi dari banyaknya *cluster* yang akan dibentuk. Metode *elbow* memiliki tujuan untuk menentukan nilai *k* terkecil dan memiliki nilai *within* yang rendah [18]. Penentuan optimasi nilai *k* ini dilihat dari nilai perbandingan dari perhitungan SSE (*Sum of Square Error*) antara jumlah klaster yang membuat suatu siku pada suatu titik. Pada gambar 3 berikut merupakan grafik dari metode *elbow* yang digunakan dalam penentuan jumlah *cluster*.



Gambar 3. Grafik hasil metode *elbow*

Cluster yang dibentuk yaitu 3 *cluster*, diantaranya *cluster* rendah (C0) yang menunjukkan sebaran UMKM rendah, *cluster* sedang (C1) menunjukkan sebaran UMKM sedang, dan *cluster* tinggi (C2) menunjukkan sebaran UMKM tinggi. Langkah awal dalam proses *clustering k-means* yaitu menentukan centroid awal yang ditentukan berdasarkan nilai terkecil untuk C0, nilai rata-rata untuk C1, dan nilai maksimum untuk C2. Atribut yang menjadi acuan dalam pengklasteran yaitu hanya atribut jumlah. Tabel 4 berikut ini merupakan *centroid* data awal pada *dataset*.

Tabel 4. *Centroid* data awal (iterasi 1)

<i>Centroid</i>	
Min (C0)	5486
Average (C1)	39285,74
Max (C2)	77167

Berdasarkan tabel 4 dapat dilihat bahwa nilai terkecil atau min (C0) berjumlah 5486, nilai rata-rata atau average (C1) berjumlah 39285,74, dan nilai maksimum atau max (C2) berjumlah 77167. Setelah menentukan *centroid* awal maka selanjutnya adalah dapat dilakukan *cluster* data dengan mengambil jarak terdekat dari tiap-tiap data. Hasil dari *clustering* yang didapatkan dengan *tools python* dapat dilihat pada gambar 4 berikut.

nama_kabupaten_kota	cluster	keterangan
KABUPATEN BOGOR	2	tinggi
KABUPATEN SUKABUMI	2	tinggi
KABUPATEN CIANJUR	0	rendah
KABUPATEN BANDUNG	2	tinggi
KABUPATEN GARUT	1	sedang
KABUPATEN TASIKMALAYA	0	rendah
KABUPATEN CIAMIS	0	rendah
KABUPATEN KUNINGAN	0	rendah
KABUPATEN CIREBON	1	sedang
KABUPATEN MAJALENGKA	1	sedang
KABUPATEN SUMEDANG	0	rendah
KABUPATEN INDRAMAYU	0	rendah
KABUPATEN SUBANG	0	rendah
KABUPATEN PURWAKARTA	1	sedang
KABUPATEN KARAWANG	1	sedang
KABUPATEN BEKASI	1	sedang
KABUPATEN BANDUNG BARAT	0	rendah
KABUPATEN PANGANDARAN	0	rendah
KOTA BOGOR	1	sedang
KOTA SUKABUMI	1	sedang
KOTA BANDUNG	1	sedang
KOTA CIREBON	1	sedang
KOTA BEKASI	1	sedang
KOTA DEPOK	1	sedang
KOTA CIMAH	1	sedang
KOTA TASIKMALAYA	1	sedang
KOTA BANJAR	1	sedang

Gambar 4. Hasil *cluster*

Pengelompokan data yang didapatkan hasil bahwa yang masuk ke dalam *cluster* rendah (C0) dengan penyebaran UMKM yang rendah diantaranya Kabupaten Cianjur, Kabupaten Tasikmalaya, Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Sumedang, Kabupaten Indramayu, Kabupaten Subang, Kabupaten Bandung Barat, dan Kabupaten Pangandaran. Untuk *cluster* sedang (C1) dengan penyebaran UMKM yang sedang diantaranya Kabupaten Garut, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar. Untuk *cluster* tinggi (C2) dengan penyebaran UMKM yang tinggi diantaranya Kabupaten Bogor, Kabupaten Sukabumi, dan Kabupaten Bandung.

3.5. Evaluation

Suatu hasil pemodelan *data mining* yang dilakukan perlu proses pengujian hasil *clustering* yang didapatkan apakah termasuk baik atau kurang baik dengan metode pengujian kualitas *cluster*. Penelitian ini akan dilakukan perbandingan dua metode pengujian yaitu *silhouette coefficient* dan *davies bouldin index* (DBI). Kriteria subjektif pengukuran pengklasteran pada *silhouette coefficient* menurut *Kauffman* dan *Roesseeuw* (1990) dapat dilihat

pada tabel 5 berikut.

Tabel 5. Kriteria evaluasi *silhouette coefficient*

Nilai SC	Kriteria
0,71 – 1,00	Struktur Kuat
0,51 – 0,70	Struktur Baik
0,26 – 0,50	Struktur Lemah
$\leq 0,25$	Struktur Buruk

Hasil penelitian ini pada pengujian *silhouette coefficient* dengan *python* pada penelitian ini didapatkan hasil sebagai berikut.

```
[ ] from sklearn.metrics import silhouette_score
```

```
[ ] silhouette_score(df2, km.labels_)
```

```
0.7348614511055411
```

Gambar 5. Hasil evaluasi *silhouette coefficient*

Pada gambar 5 di atas dapat dilihat bahwa hasil evaluasi menggunakan *silhouette coefficient* didapatkan nilai index sebesar 0,73 yang berarti masuk ke dalam kriteria pengklasteran dengan struktur yang kuat atau klaster terbaik.

```
[ ] from sklearn.metrics import davies_bouldin_score
```

```
[ ] davies_bouldin_score(df2, km.labels_)
```

```
0.29084239471768747
```

Gambar 6. Hasil evaluasi *davies bouldin index* (dbi)

Pada gambar 6 di atas menunjukkan hasil evaluasi dengan menggunakan DBI yang di mana didapatkan nilai sebesar 0,29. Di mana pada evaluasi DBI jika nilai semakin mendekati nol maka *cluster* yang dibentuk dapat dikatakan baik. Sehingga hal tersebut menyatakan bahwa evaluasi klaster *k-means* dengan DBI ini termasuk ke dalam kategori baik.

Perbandingan kedua metode evaluasi penerapan *clustering k-means* dihasilkan bahwa menggunakan evaluasi metode *silhouette coefficient* dan metode *davies bouldin index* keduanya menunjukkan hasil evaluasi yang baik pada kasus *clustering* persebaran UMKM di Jawa Barat.

4. KESIMPULAN

Hasil penelitian ini menunjukkan *clustering k-means* dapat diimplementasikan dalam pengklasteran UMKM di Jawa Barat. Didapat 3 *cluster* yang dibentuk, di mana yang masuk ke dalam *cluster* 0 atau tingkat penyebaran UMKM rendah berjumlah 9 Kabupaten/Kota, yang masuk ke dalam *cluster* 1 atau tingkat penyebaran UMKM sedang berjumlah 15 Kabupaten/Kota, dan yang masuk ke dalam *cluster* 2 atau tingkat penyebaran UMKM tinggi berjumlah 3 Kabupaten/Kota. Diharapkan Kabupaten/Kota tersebut mendapatkan perhatian pemerintah atau institusi terkait dalam penetapan strategi persebaran dan peningkatan jumlah UMKM di Jawa Barat. Pada penelitian ini hasil evaluasi *clustering* dengan menerapkan algoritma *k-means* yang dilakukan dengan membandingkan dua metode evaluasi menggunakan *silhouette coefficient* yang di mana didapatkan nilai sebesar 0,73 yang berarti masuk ke dalam kriteria klaster terbaik. Sedangkan dengan menggunakan metode *davies bouldin index* (DBI) didapatkan nilai sebesar 0,29 yang berarti masuk ke dalam kategori baik. Dari hasil perbandingan kedua evaluasi *cluster*

tersebut menunjukkan hasil *cluster* dengan menerapkan algoritma *k-means* yang terbentuk baik.

5. SARAN

Pengimplementasian algoritma *k-means* ini sudah menunjukkan hasil yang cukup baik untuk pengelompokan data ke dalam *cluster*. Penelitian ini perlu dilakukan beberapa metode evaluasi lainnya guna mengetahui apakah masih menghasilkan hasil cukup baik atau tidak.

DAFTAR PUSTAKA

- [1] www.bi.go.id (2021, 03 Juni). Laporan Perekonomian Provinsi Jawa Barat Juni 2021. Diakses pada 01 November 2021, dari <https://www.bi.go.id/id/publikasi/laporan/lpp/Pages/Laporan-Perekonomian-Provinsi-Jawa-Barat-Juni-2021.aspx>
- [2] Puntoriza, P., dan Fibriani, C., 2020, Analisis Persebaran UMKM Kota Malang Menggunakan Cluster K-means, JOINS (Journal of Information System), 5(1), 86–94. <https://doi.org/10.33633/joins.v5i1.3469>.
- [3] Idayu, R., Husni, M., dan Suhandi, S., 2021, Strategi Pengembangan Usaha Mikro Kecil dan Menengah (UMKM) untuk Meningkatkan Perekonomian Masyarakat Desa Di Desa Nembol Kecamatan Mandalawangi Kabupaten Pandeglang Banten, Jurnal Manajemen STIE Muhammadiyah Palopo, 7(1), 73. <https://doi.org/10.35906/jm001.v7i1.729>.
- [4] Siska, S. T., 2016, Analisa dan Penerapan Data Mining untuk Menentukan Kubikasi Air Terjual Berdasarkan Pengelompokan Pelanggan Menggunakan Algoritma K-Means Clustering, Jurnal Teknologi Informasi & Pendidikan, 9(1), 86–93.
- [5] Astuti, D., Iskandar, A. R., dan Febrianti, A., 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>.
- [6] Aristika, W., dan Hartono, W. J., 2020, Penerapan Clustering K-Means untuk Menentukan Pengaruh Media Sosial Facebook Terhadap Usaha Mikro, Kecil dan Menengah (UMKM) Di Kecamatan Pekanbaru Kota, Jurnal Ilmu Komputer dan Bisnis, 11, 2389–2395.
- [7] Fitriani, dan Elisa, E., 2021, Pemanfaatan Data Mining Clustering dalam Penentuan Media Promosi UMKM Di Kota Batam, Computer and Science Industrial Engineering (COMASIE), 4(1), 59–65.
- [8] Setiawan, R., 2016, Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik LP3I Jakarta), Jurnal Lentera Ict, 3(1), 76–92.
- [9] Maulana, A., dan Fajrin, A. A., 2018, Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor, Klik - Kumpulan Jurnal Ilmu Komputer, 5(1), 27, <https://doi.org/10.20527/klik.v5i1.100>.
- [10] Fitriyyah, S. N. J., Safriadi, N., dan Pratama, E. E., 2019, Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive bayes. Jurnal Edukasi dan Penelitian Informatika (JEPIN), 5(3), 279, <https://doi.org/10.26418/jp.v5i3.34368>.
- [11] Putra, R. R., dan Wadisman, C., 2018, Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means, Intecom: Journal of Information

- Technology and Computer Science, 1(June), 72–77, <https://doi.org/https://doi.org/10.31539/intecom.v1i1.141>.
- [12] Gustientiedina, Adiya, M. H., dan Desnelita, Y., 2019, Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru, *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24, <https://doi.org/10.25077/teknosi.v5i1.2019.17-24>.
- [13] Fahmi, R. N., Jajuli, M., dan Sulistiyowati, N., 2021, Analisis Pemetaan Tingkat Kriminalitas Di Kabupaten Karawang Menggunakan Algoritma K-Means, *Journal Of Information Technology And Computer Science (INTECOMS)*, 4, 67–79.
- [14] Ramadhani, R. D., dan Ak, D. J., 2017, Evaluasi K-Means dan K-Medoids pada Dataset Kecil, *Seminar Nasional Informatika dan Aplikasinya*, September, 20–24, [file:///C:/Users/User/Downloads/Evaluasi K-Means dan K-Medoids pada Dataset Kecil.pdf](file:///C:/Users/User/Downloads/Evaluasi%20K-Means%20dan%20K-Medoids%20pada%20Dataset%20Kecil.pdf).
- [15] Gie, W., dan Jollyta, D., 2020, Perbandingan Euclidean dan Manhattan untuk Optimasi Cluster Menggunakan Davies Bouldin Index : Status Covid-19 Wilayah Riau, *Prosiding Seminar Nasional Riset dan Information Science (SENARIS) 2020*, 2(April), 187–191.
- [16] Jollyta, D., Efendi, S., Zarlis, M., dan Mawengkang, H., 2019, Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index, *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(September), 918, <https://doi.org/10.30645/senaris.v1i0.100>.
- [17] Febrinanto, F. G., Dewi, C., dan Wiratno, A. T., 2018, Implementasi Algoritme K-Means Sebagai Metode Segmentasi Citra Dalam Identifikasi Penyakit Daun Jeruk, *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(11), 5375–5383.
- [18] Hartanti, N. T., 2020, Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional, *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(2), 82–89, <https://doi.org/10.25077/teknosi.v6i2.2020.82-89>.
-