

## Pemetaan UMKM dalam Upaya Pengentasan Kemiskinan dan Penyerapan Tenaga Kerja Menggunakan Algoritma K-Means

Herwinda Kurniadewi<sup>1\*</sup>, Rijal Abdul Hakim<sup>2\*</sup>, Mohamad Jajuli<sup>3\*</sup>, Jajam Haerul Jaman<sup>4\*</sup>

\* Teknik Informatika, Universitas Singaperbangsa Karawang

[herwinda.kurniadewi18188@student.unsika.ac.id](mailto:herwinda.kurniadewi18188@student.unsika.ac.id)<sup>1</sup>, [rijal.16182@student.unsika.ac.id](mailto:rijal.16182@student.unsika.ac.id)<sup>2</sup>, [mohamad.jajuli@unsika.ac.id](mailto:mohamad.jajuli@unsika.ac.id)<sup>3</sup>, [jajam.haeruljaman@staff.unsika.ac.id](mailto:jajam.haeruljaman@staff.unsika.ac.id)<sup>4</sup>

### Article Info

#### Article history:

Received 2022-07-04

Revised 2022-08-20

Accepted 2022-08-27

#### Keyword:

Algoritma K-Means,  
Clustering,  
Python,  
Silhouette coefficient.

### ABSTRACT

Covid pandemic created an economic crisis. Increase the poverty rate by double digits in one year in Indonesia. Covid pandemic has also had an impact on Indonesia's employment conditions, such as finding it difficult to find work. Absorption of labor has a close correlation with poverty. The workforce has a significant influence on the poverty level. One of the regencies in West Java which has a high poverty rate and job seekers is increasing compared to the previous year, Purwakarta Regency. Poverty alleviation by developing MSMEs has good potential. The development of MSMEs will be able to absorb more workers and increase people's income so that it can encourage the rate of economic growth. In this study using the CRISP-DM methodology. In this study, MSMEs in Purwakarta Regency were grouped based on location, number of MSMEs, number of poor people and number of job seekers by using the k-means algorithm and mapping using python. The results of the grouping obtained 3 clusters, namely clusters as many as 6 districts, clusters as many as 8 districts and clusters as many as 3 districts. To determine the performance of the model, an evaluation of the silhouette coefficient which obtained a value of 0.45.



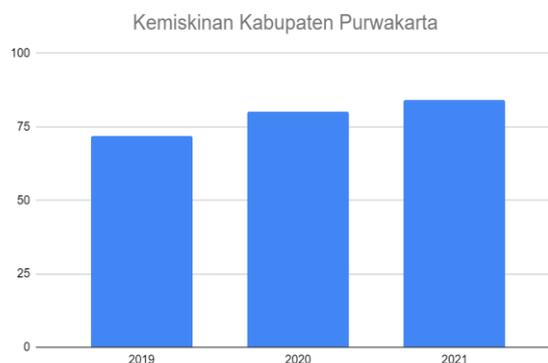
This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

### I. PENDAHULUAN

Pada Tahun 2020 telah terjadi Pandemi Covid 19 yang menciptakan krisis ekonomi dimana salah satu parameternya yaitu meningkatnya angka kemiskinan [1]. Kemiskinan dapat diartikan yaitu jika sebuah rumah tangga tidak mempunyai cukup kemampuan untuk memenuhi kebutuhan hidup. Kemiskinan adalah perkara multi dimensi dan juga merupakan permasalahan umum yang dihadapi negara berkembang [2]. Warga miskin di Indonesia mencapai 74 juta penduduk. Jumlah ini bertambah per Maret 2021 sebanyak 37,54 juta warga miskin sesuai dengan data yang telah diverifikasi Kemensos [3]. Jawa Barat, Jawa Timur, dan Jawa Tengah adalah provinsi dengan jumlah penduduk miskin ekstrem pada Maret 2021, meskipun memiliki persentase penduduk miskin ekstrem yang rendah. Jumlah penduduk ekstrem di Jawa Barat menyentuh angka 1,8 juta orang dan persentase dengan nilai 3,6% dari jumlah penduduknya[4].

Salah satu Kabupaten di Jawa Barat yang memiliki angka kemiskinan yang cukup tinggi adalah Kabupaten Purwakarta atau dikenal sebagai kota Istimewa. Angka kemiskinan di

Kabupaten Purwakarta Tahun 2020 meningkat dibanding tahun sebelumnya. Jumlah masyarakat masuk kategori miskin di Kabupaten Purwakarta per Maret 2019 yaitu 71.86 ribu orang, terus menaik hingga tahun 2021 meningkat kembali menjadi 84.27 ribu orang.



Gambar 1. Kemiskinan Kabupaten Purwakarta

Pandemi Covid-19 juga berdampak terhadap kondisi ketenagakerjaan Indonesia seperti mulai sulit mencari pekerjaan. Pencarian kerja di tengah pandemi mengalami peningkatan hingga tahun 2021. Persaingan ini lebih sulit lagi karena adanya angkatan kerja baru dan pekerja yang kehilangan posisi di perusahaan akibat terkena dampak pandemi [5][6]. Angka pelamar kerja juga turut mengalami kenaikan sebesar 37% dibandingkan tahun lalu [7]. Akan tetapi, lowongan kerja terus menurun [8].

Ancaman utama untuk angkatan kerja baru dalam masa pandemi ini juga adalah semakin sempitnya lapangan kerja. Pekerja yang dirumahkan turut menyumbang peningkatan angka pengangguran di Purwakarta. Selain itu terjadi penurunan serapan tenaga kerja sektor industri pengolahan di masa pandemi Covid-19 [9] [10].

Penyerapan tenaga kerja memiliki kolerasi yang erat dengan kemiskinan. Tenaga kerja mempunyai pengaruh yang signifikan terhadap tingkat kemiskinan [11], dengan terus adanya peningkatan penyerapan tenaga kerja maka akan mengakibatkan jumlah kemiskinan yang menurun [12][13].

Usaha dalam menanggulangi pengentasan kemiskinan ini harus melihat akar permasalahan kemiskinan, salah satunya yaitu tidak adanya pekerjaan yang menyebabkan rendahnya pendapatan penduduk miskin. Upaya penanggulangan kemiskinan dan pengangguran (ketenagakerjaan) juga merupakan salah satu pokok strategis dalam Rencana Pembangunan Jangka Menengah Daerah (RPJMD) Kabupaten Purwakarta 2018-2023.

Pengentasan kemiskinan melalui pengembangan UMKM memiliki potensi yang cukup baik karena dapat menyerap lebih banyak tenaga kerja dan meningkatkan pendapatan masyarakat sehingga dapat mendorong laju pertumbuhan ekonomi [14]. Dinas Tenaga Kerja dan Transmigrasi Purwakarta juga mencoba untuk memulihkan perekonomian warga melalui penciptaan wirausaha baru agar dapat menyerap lebih banyak tenaga kerja.

Pelaku utama ekonomi di Indonesia dikuasai oleh Usaha Mikro, Kecil dan Menengah (UMKM). Dalam UUD 1999 Pasal 33 ayat 4, UMKM adalah elemen dari perekonomian nasional yang berwawasan kemandirian dan mempunyai peluang yang besar guna mencapainya peningkatan kesejahteraan masyarakat. Peran UMKM dalam pertumbuhan ekonomi sangat signifikan. UMKM di Indonesia hingga tahun 2012 mampu mendapatkan 85-107 juta tenaga kerja [15]. UMKM memiliki peran dalam pertumbuhan pembangunan dan ekonomi. Pengelompokan UMKM di Kabupaten Purwakarta harus dilakukan untuk mempermudah pemerintah dalam upaya pengembangan dan pemberdayaan UMKM yang terdapat di Kabupaten Purwakarta agar bisa tepat sasaran. Selain itu juga dapat membantu pemerintah dalam hal terkait peminjaman modal, penentuan peluang usaha dan penerapan rencana pemasaran yang sesuai untuk mengembangkan pasar sebagai prioritas utama.

*Clustering* adalah salah satu bidang penelitian dalam data mining dan mempunyai fungsi untuk pengelompokan data atau pemetaan data berdasarkan keunikannya [16].

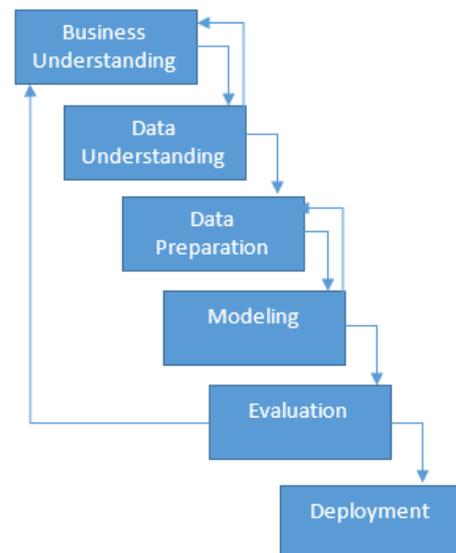
Algoritma K-Means merupakan metode *clustering* data mining dengan proses modelling tanpa pengawasan. Data dikelompokkan dengan metode k-means kedalam beberapa golongan dan masing-masing golongan mempunyai keunikan yang serupa dengan lainnya tetapi berbeda dari golongan yang lain. Tujuannya adalah meminimalkan perbedaan masing-masing data di dalam *cluster* dan memaksimalkan perbedaan cluster lainnya [17] [18].

Pada penelitian analisis *clustering* k-means dengan memakai data kemiskinan di Jawa Barat melakukan pemetaan karakteristik berdasarkan rata-rata nilai yang paling tinggi dan nilai yang paling rendah dari setiap parameter kemiskinan dan mendapatkan hasil 5 *cluster* [19]. Pada tahun selanjutnya dilakukan analisis persebaran UMKM kota Malang dengan *cluster* k-means menggunakan data dari Dinas Koperasi dan UMKM Kota Malang [20].

Pada Penelitian ini akan dilakukan pemetaan UMKM di Kabupaten Purwakarta berdasarkan lokasi/kecamatan, jumlah UMKM, jumlah penduduk miskin dan jumlah pencari kerja. Adapun hasil yang diharapkan yaitu berupa visualisasi pemetaan yang memuat hasil *cluster* pada setiap lokasi/kecamatan. Diharapkan pemetaan ini dapat membantu pemerintah dalam pengentasan kemiskinan.

## II. METODOLOGI PENELITIAN

Penelitian ini menggunakan metodologi CRISP-DM. Metode CRISP-DM (*Cross Industry Standard for Data Mining*) adalah standarisasi teknik pengelolaan data dan cukup umum digunakan oleh ahli untuk pemecahan kasus [21]. CRISP-DM diciptakan oleh Daimler Chrysler, SPSS, NCR sebagai standarisasi *data mining*. Kemudian dikembangkan pada berbagai workshop [22]. Penelitian ini menggunakan algoritma K-Means dengan metode *clustering*. Penelitian ini akan dilakukan dalam 6 tahapan sebagai berikut.



Gambar 2. Tahapan Penelitian

#### A. Business Understanding

Fase ini merupakan fase awal dari model CRISP-DM. Dalam fase ini akan dilakukan studi pustaka mengenai objek penelitian yang akan diteliti. Beberapa kasus *clustering* dan pemetaan yang berkaitan dengan UMKM, kemiskinan ataupun pengangguran juga akan dikaji sehingga solusi yang ditawarkan diharapkan menjadi efektif. Proses data mining ditawarkan sebagai solusi dengan tujuan untuk menyelesaikan permasalahan yang telah dikemukakan. Beberapa kriteria keberhasilan juga ditentukan dalam di tahap ini.

#### B. Data Understanding

Pada fase kedua yaitu *data understanding*. Pada fase ini dilakukan pengumpulan data awal sebagai tolak ukur dari kualitas data. Data tersebut akan dipelajari dan dipahami lebih lanjut. Selain itu, pada tahap ini juga dilakukan deskripsi data termasuk atribut, tipe data dan keterangan atribut.

#### C. Data Preparation

Pada tahap ini dilakukan preprocessing, yaitu fase mempersiapkan data yang digunakan sebelum memasuki tahap *modeling*. Pada fase ini terdapat beberapa tahapan:

- 1) *Data selection*, proses ini mengidentifikasi atribut yang akan dipakai dan menghapus atribut yang tidak mempunyai korelasi dengan penelitian ini.
- 2) *Data preprocessing*, proses ini mengidentifikasi dan membersihkan *missing values* dan memeriksa data yang tidak konsisten.
- 3) *Transformation*, yaitu proses pengelompokan atribut ke dalam data yang baru, lalu mengintegrasikan data hasil *data preprocessing*, dan mentransformasikan data untuk kemudian diproses lebih lanjut

#### D. Modeling

Fase *modeling* bertujuan untuk memilih model dan teknik *data mining* yang sesuai dan akan digunakan. Pada fase ini diterapkan metode *clustering* dengan algoritma yang akan digunakan yaitu algoritma k-means.

Algoritma K-means adalah salah satu bentuk pengelompokan data yang sederhana. K-Means adalah algoritma yang dipakai untuk *clustering* dan berjenis *non-hierarchical* dengan mempunyai waktu komputasi cukup cepat [23]. K-Means merupakan metode *clustering data mining* dengan proses modeling tanpa pengawasan dan mengelompokkan data secara partisi. Data dikelompokkan dengan metode k-means kedalam beberapa golongan dan masing-masing golongan mempunyai keunikan yang serupa dengan lainnya tetapi berbeda dari golongan yang lain. Tujuannya adalah meminimalkan perbedaan masing-masing data di dalam *cluster* dan memaksimalkan perbedaan *cluster* lainnya [17].

#### E. Evaluation

Pada fase *evaluation* akan menganalisis hasil pemodelan *data mining* yang digunakan. Metode evaluasi yang digunakan yaitu metode *silhouette coefficient*. *Silhouette Coefficient* adalah teknik dengan tujuan untuk melihat

kualitas *cluster*. Metode ini adalah metode kombinasi dari metode cohesion dan metode separation. Fungsi dari metode ini adalah digunakan untuk menghitung dekatnya objek-objek suatu cluster, dan metode separation yang mempunyai fungsi menghitung jarak sebuah *cluster* terhadap *cluster* lainnya [24].

Metode ini digunakan untuk melakukan uji kualitas dari model. Pada fase *evaluation* ini dilakukan untuk dapat menilai apakah model yang dibangun sudah sesuai dengan tujuan yang telah ditentukan. Jika hasil yang didapatkan sudah sesuai dengan kriteria, maka dapat dilakukan fase terakhir.

#### F. Deployment

Pada fase terakhir yaitu fase *deployment*. Fase ini merupakan tahap penyebaran terhadap hasil dari seluruh fase yang telah dilakukan dalam penelitian ini. Pada fase *deployment* dilakukan dengan membuat laporan presentasi dari pengetahuan yang didapatkan berdasarkan hasil pemodelan dan evaluasi. Hasil yang didapatkan dapat digunakan pemerintah dalam upaya pengembangan dan pembinaan UMKM di Kabupaten Purwakarta.

### III. HASIL DAN PEMBAHASAN

Hasil dari penelitian yang telah dilakukan adalah melakukan pengelompokan UMKM dalam Upaya Pengentasan Kemiskinan dan Penyerapan Tenaga Kerja di Kabupaten Purwakarta dengan menggunakan algoritma k-means *clustering*. Hasil pengelompokan mendapatkan 3 cluster yaitu prioritas tinggi, prioritas rendah, dan prioritas sedang. Selanjutnya hasil dari pengelompokan tersebut dievaluasi menggunakan *silhouette coefficient* dengan library sklearn dan divisualisasikan menggunakan library matplotlib.pyplot dan geopandas. Seluruh tahapan menggunakan bahasa pemrograman python.

#### A. Business Understanding

Pada tahapan *business understanding* dilakukan pemahaman tujuan kebutuhan berdasarkan penilaian bisnis. Kemudian diubah menjadi sebuah rencana awal yang dirancang untuk mencapai tujuan Berikut langkah dalam pemahaman bisnis, yaitu

##### 1) Determine Business Objectives

Dalam penelitian ini, yang menjadi target pengguna yaitu Pemerintah Daerah Kabupaten Purwakarta. Adapun dari perspektif bisnis, yang dibutuhkan oleh pengguna untuk membantu dalam pengentasan kemiskinan melalui UMKM adalah pemetaan UMKM berdasarkan jumlah penduduk miskin dan jumlah pencari kerja tahun 2021 di Kabupaten Purwakarta.

##### 2) Assess Situation

Adapun kondisi saat ini dilihat dari fakta di lapangan menunjukkan bahwa Instansi-instansi Pemerintah Daerah Kabupaten Purwakarta terkait belum memiliki visualisasi pemetaan wilayah yang berdasarkan mengenai kemiskinan dan UMKM yang informatif dan dikhususkan dalam permasalahan ini. Adapun kondisi data pada Dinas UMKM yaitu memiliki data lengkap mengenai data UMKM untuk

semua kecamatan di Kabupaten Purwakarta. Sedangkan data-data penduduk miskin dan pencari kerja bersifat rahasia dan hanya didapatkan atribut-atribut yang dibutuhkan.

### 3) *Determine Data Mining Goals*

Tujuan data mining yang dilakukan dalam penelitian ini yaitu melakukan pengklasteran UMKM dengan menggunakan metode *clustering* dengan menerapkan algoritma k-means yang akan divisualisasikan kemudian untuk mempermudah pemerintah dalam mengetahui penyebaran penduduk miskin, pencari kerja dan UMKM.

### 4) *Plan Activities*

Guna mencapai tujuan data mining dalam pengklasteran UMKM di Kabupaten Purwakarta, tools yang akan digunakan yaitu menggunakan jupyter notebook dengan bahasa pemrograman python.

## B. *Data Understanding*

Fase berikutnya adalah fase pemahaman data awal, mengenal, dan memahami data yang dimiliki serta melakukan analisis yang bisa dilakukan pada data tersebut. Data yang digunakan dalam penelitian adalah Data UMKM yang bersumber dari Dinas Koperasi UKM, Perdagangan, dan Perindustrian di Kabupaten Purwakarta, Data Penduduk Miskin Tahun yang bersumber dari Dinas Sosial Pemberdayaan Perempuan dan Perlindungan Anak di Kabupaten Purwakarta, dan Data Pencari Kerja yang bersumber dari Dinas Ketenagakerjaan dan Transmigrasi di Kabupaten Purwakarta. Data yang digunakan untuk dianalisis yaitu data pada tahun 2021. Contoh Data awal dapat dilihat pada Tabel 1, Tabel 2 dan Tabel 3 berikut:

TABEL 1  
DATA UMKM PURWAKARTA

| No | Nama                | Bidang Usaha                  | Alamat  |
|----|---------------------|-------------------------------|---|
| 1  | Sukaenah            | Jual Bubur                    | Purnawarman Selatan Rt 42 Rw 08 Kel. Sindangkasih Kec. Purwakarta |
| 2  | Asep Saripudin      | Wrung Sembako                 | Gg. Aster Iii No. 118 Rt009 Rw002 Kel. Nagrikaler Kec. Purwakarta |
| 3  | Ujang Hadibyoy      | Jualan Tempe Mendoan          | Gg Aster I Rt 12 Rw 02 Kel. Nagrikaler Kec. Purwakarta            |
| 4  | Oktavianis Safarina | Jualan Sosis                  | Bongas Rt 34 Rw 04 Kel. Sindangkasih Kec. Purwakarta              |
| 5  | Sri Hasanah         | Krediton Perabot Rumah Tangga | Purnawarman Selatan Rt 42 Rw 08 Kel. Sindangkasih Kec. Purwakarta |

TABEL 2  
DATA PENDUDUK MISKIN

| No | Kecamatan    | Desa/Kelurahan | Jenis Kelamin | Alamat      |
|----|--------------|----------------|---------------|-------------|
| 1. | Babakancikao | Babakancikao   | L             | Kp. Karajan |
| 2. | Babakancikao | Babakancikao   | L             | Kp. Tabrik  |
| 3. | Babakancikao | Babakancikao   | P             | Kp. Karajan |
| 4. | Babakancikao | Babakancikao   | P             | Kp. Tabrik  |
| 5. | Babakancikao | Babakancikao   | P             | Kp. Karajan |

TABEL 3  
DATA PENCARI KERJA PURWAKARTA

| No | Kabupaten  | Kecamatan  | Jenis kelamin | Tingkat Pendidikan |
|----|------------|------------|---------------|--------------------|
| 1  | Purwakarta | Purwakarta | L             | SD                 |
| 2  | Purwakarta | Purwakarta | L             | SD                 |
| 3  | Purwakarta | Purwakarta | L             | SD                 |
| 4  | Purwakarta | Purwakarta | L             | SD                 |
| 5  | Purwakarta | Purwakarta | L             | SD                 |

Untuk memahami data yang akan diolah dalam penelitian ini lebih dalam, dilakukan Describe the data. Dalam tahap ini data akan dideskripsikan mengenai atribut, tipe data, dan keterangan data. Data ini memiliki beberapa atribut dengan penjelasan atribut seperti pada Tabel 4, Tabel 5 dan Tabel 6.

TABEL 4  
DESKRIPSI DATA UMKM PURWAKARTA

| Atribut      | Tipe Data | Keterangan                                  |
|--------------|-----------|---|
| No           | Nominal   | No urut daftar UMKM di Kabupaten Purwakarta |
| Nama         | Nominal   | Nama Pemilik Usaha                          |
| Bidang Usaha | Nominal   | Jenis Usaha                                 |
| Alamat       | Nominal   | Alamat Usaha                                |

TABEL 5  
DESKRIPSI DATA PENDUDUK MISKIN PURWAKARTA

| Atribut         | Tipe Data | Keterangan                                  |
|-----------------|-----------|---|
| No              | Nominal   | No urut penduduk miskin                     |
| Kecamatan       | Nominal   | Nama Kecamatan di Kabupaten Purwakarta      |
| Desa/ Kelurahan | Nominal   | Nama Desa/Kelurahan di Kabupaten Purwakarta |
| Jenis Kelamin   | Nominal   | Jenis Kelamin Penduduk                      |
| Alamat          | Nominal   | Alamat Penduduk                             |

TABEL 6  
DESKRIPSI DATA PENCAIRI KERJA

| Atribut            | Tipe Data | Keterangan                             |
|--------------------|-----------|--|
| No                 | Nominal   | No urut pencari kerja                  |
| Kecamatan          | Nominal   | Nama Kecamatan di Kabupaten Purwakarta |
| Jenis Kelamin      | Nominal   | Jenis Kelamin Penduduk                 |
| Tingkat Pendidikan | Nominal   | Tingkat Pendidikan                     |

### C. Data Preparation

Pada tahap data preparation dilakukan kegiatan persiapan data dengan membangun suatu data untuk menyesuaikan dengan kebutuhan pada pemodelan dari data mentah.

#### 1) Data Selection

Data yang diambil yaitu data UMKM, penduduk miskin, dan pencari kerja di Kabupaten Purwakarta. Data ini disimpan dalam bentuk excel. Lalu dibersihkan dengan menghapus atribut-atribut yang tidak diperlukan pada penelitian ini. Data yang dipilih yaitu pada tahun 2021. Dalam data umkm akan dilakukan penambahan atribut kecamatan pada data UMKM. Proses penambahan atribut kecamatan ini dilakukan mengacu kepada atribut alamat. Dalam data penduduk miskin dilakukan penghapusan kolom Desa/kelurahan. Sedangkan pada data pencari kerja akan dilakukan penghapusan kolom kabupaten.

#### 2) Data Preprocessing

Berdasarkan tahap *data understanding* dan verifikasi kualitas data menunjukkan bahwa adanya data dengan kualitas tidak baik dan perlu diperbaiki. Kemudian langkah selanjutnya yaitu menghapus *missing value* atau nilai kosong pada ketiga data.

```
data1.isna().sum()
No                0
NAMA              0
BIDANG USAHA     0
ALAMAT           0
KECAMATAN        0
dtype: int64

Data2.isna().sum()
No                0
KECAMATAN        0
JENIS KELAMIN    0
ALAMAT           0
dtype: int64

Data3.isna().sum()
No                0
KECAMATAN        0
JENIS KELAMIN    0
TINGKAT PENDIDIKAN 0
```

Selanjutnya akan dilakukan pengelompokkan data. Data akan dikelompokkan berdasarkan kecamatan guna mendapatkan atribut yang akan digunakan untuk pemodelan yaitu jumlah UMKM, jumlah penduduk miskin, jumlah

pencari kerja menggunakan fungsi *agregasi count* sesuai pada tahap *business understanding* dan memastikan data sudah siap untuk digunakan pada tahap *modeling*.

TABEL 7  
DATA INTEGRATION

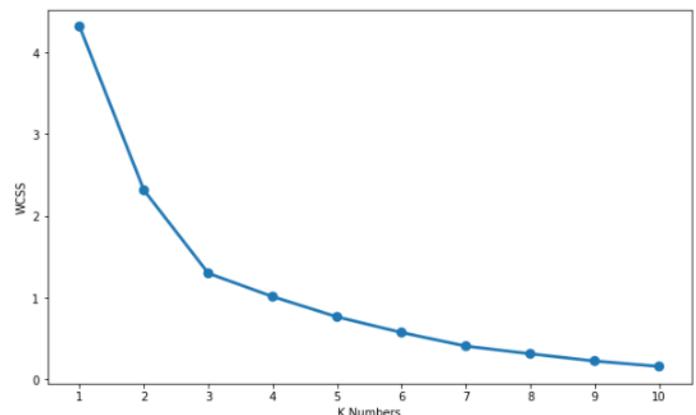
| No | Kecamatan      | Jumlah UMKM | Jumlah Penduduk Miskin | Jumlah Pencari Kerja |
|----|----------------|-------------|------------------------|----------------------|
| 1. | Bababakancikao | 163         | 19433                  | 1481                 |
| 2. | Bojong         | 102         | 37369                  | 716                  |
| 3. | Bungursari     | 299         | 19710                  | 1447                 |
| 4. | Campaka        | 335         | 21239                  | 1998                 |
| 5. | Cibatu         | 342         | 17294                  | 1297                 |

#### 3) Transformation

Data yang digunakan dalam penelitian ini mempunyai tipe data kategorik dan numerik. Sedangkan dalam menggunakan pemodelan data yang digunakan harus memiliki tipe data numerik atau angka. Dalam data ini yang memiliki tipe data kategorik yaitu atribut kecamatan. Oleh sebab itu data harus ditransformasikan dari kategorik menjadi numerik menggunakan *label encoder*. Selanjutnya adalah melakukan penanganan data numerik agar mempunyai rentang nilai yang sama dan agar tidak terdapat data yang terlalu besar sampai terlalu kecil dan juga melakukan normalisasi data menggunakan *min-max scaler* dan *normalize*.

### D. Modeling

Tahap selanjutnya adalah *modeling* yang merupakan fase penerapan *data mining* yang digunakan sesuai dengan tujuan. Pada penelitian ini teknik *data mining* yang digunakan yaitu clustering dengan menggunakan algoritma K-Means. Tahap pertama dalam fase ini yaitu menentukan jumlah *cluster*. Untuk menentukan jumlah cluster akan menggunakan metode elbow seperti pada Gambar 3 terlihat garis mengalami patahan yang membentuk elbow atau siku pada saat  $K = 3$ . Setelah mengetahui  $K$  optimal dilakukan proses *clustering* dengan algoritma k-means. Seperti pada Gambar 4.



Gambar 3. Metode Elbow

```
kmeans = KMeans(n_clusters =3, init='k-means++',
                random_state=42)
model = kmeans.fit(normalized_df)
label = kmeans.predict(normalized_df)
```

TABEL 8  
HASIL CLUSTERING DATA PENDUDUK PURWAKARTA

| Kecamatan    | Jumlah UMKM | Jumlah Penduduk Miskin | Jumlah Pencari Kerja | cluster |
|--------------|-------------|------------------------|----------------------|---------|
| Babakancikao | 163         | 19433                  | 1481                 | 1       |
| Bojong       | 102         | 37369                  | 716                  | 0       |
| Bungursari   | 299         | 19710                  | 1447                 | 1       |
| Campaka      | 335         | 21239                  | 1998                 | 1       |
| Cibatu       | 342         | 17294                  | 1297                 | 1       |
| Darangdan    | 210         | 43782                  | 1130                 | 0       |
| Jatiluhur    | 257         | 33890                  | 2013                 | 1       |
| Kiarapedes   | 361         | 20740                  | 452                  | 2       |
| Maniis       | 110         | 29847                  | 415                  | 2       |
| Pasawahan    | 171         | 28829                  | 1477                 | 2       |
| Plered       | 187         | 59090                  | 1477                 | 0       |
| Pondok Salam | 82          | 21222                  | 818                  | 2       |
| Purwakarta   | 861         | 54088                  | 4448                 | 1       |
| Sukasari     | 70          | 12171                  | 178                  | 2       |
| Sukatani     | 97          | 41156                  | 1452                 | 2       |
| Tegal Waru   | 122         | 39274                  | 653                  | 2       |
| Wanayasa     | 210         | 29970                  | 702                  | 2       |

Pada Tabel 8 diatas menunjukkan hasil clustering data penduduk Purwakarta, dimana cluster 0 terdapat 3 anggota yaitu bojong,arangdan, plered, cluster 1 terdapat 6 anggota yaitu babakancikao, bungursari, campaka, cibatu, jatiluhur, purwakarta, dan cluster 2 terdapat 8 anggota yaitu kiarapedes, maniis, pasawahan, pondok salam, sukasari, sukatani, tegal waru, dan wanayasa. Kemudian untuk membedakan hasil pengelompokan yang terbentuk maka dilakukan profilisasi dengan mencari nilai rata-rata dari setiap variabel. Penamaan cluster akan ditentukan dari atribut yang digunakan dalam *modeling* yaitu jumlah pencari kerja lalu jumlah penduduk miskin dan jumlah UMKM.

| clusters | Jumlah UMKM | Jumlah Penduduk Miskin | Jumlah Pencari Kerja |
|----------|-------------|------------------------|----------------------|
| 0        | 166.333333  | 46747.000              | 1107.666667          |
| 1        | 376.166667  | 27609.000              | 2114.000000          |
| 2        | 152.875000  | 27901.125              | 768.375000           |

Gambar 4. Karakteristik cluster

Pada Gambar 4 dapat diketahui karakteristik dari setiap cluster yang terbentuk. Pada Cluster 0 terdapat rata-rata dengan jumlah pencari kerja sebesar 1107.67, jumlah penduduk miskin 4647.000, dan jumlah umkm sebesar 166.34. Cluster 0 memiliki karakteristik dengan jumlah pencari kerja sedang, jumlah penduduk miskin tinggi dan jumlah umkm sedang. Berdasarkan hal tersebut, maka cluster 0 sebagai prioritas sedang.

Pada cluster 1 memiliki jumlah pencari kerja sebesar 2114.00, jumlah penduduk miskin 27609.000 dan rata-rata dengan jumlah umkm sebesar 376,167. Pada cluster 1 memiliki karakteristik dengan jumlah pencari kerja tinggi, jumlah penduduk miskin rendah, dan jumlah UMKM tinggi. Berdasarkan hal tersebut, maka cluster 1 sebagai prioritas tinggi. Sedangkan pada cluster 2 memiliki rata-rata jumlah pencari kerja sebesar 768.375, jumlah penduduk miskin 27901.125 dan jumlah umkm sebesar 152.875. Pada cluster 2 memiliki karakteristik dengan jumlah pencari kerja rendah, jumlah penduduk miskin rendah, dan jumlah UMKM rendah. Berdasarkan hal tersebut, maka cluster 2 sebagai prioritas rendah.

E. Evaluation

Setelah melakukan tahapan *modeling* maka perlu dilakukan evaluasi guna mengetahui tingkat keberhasilan pemodelan yang diterapkan apakah sudah sesuai dengan tujuan yang sudah dirancang pada tahap *business understanding*.

1) Evaluate Result

Hasil clustering dengan algoritma k-means akan dilakukan menggunakan *silhouette coefficient* dengan library sklearn. Hasil evaluasi performa yang didapatkan dengan *cluster* terbaik adalah pada 3 *cluster* yaitu 0.45 yang mendekati nilai 1.

```
from sklearn.metrics import silhouette_score
silhouette_score(normalized_df, label)

0.45117104577271494
```

Gambar 5. Evaluasi silhouette coefficient

2) Determine Next Step

Pada tahapan ini dilakukan penentuan langkah selanjutnya yang akan diambil. Terdapat dua pilihan yaitu kembali ke tahap *business understanding* atau melanjutkan ke tahap *deployment*. Dari hasil evaluasi menunjukkan bahwa melakukan penyelesaian tahap akhir proses CRISP-DM yaitu *deployment*.

F. Deployment

Setelah proses pemodelan dengan menerapkan teknik *data mining clustering*, selanjutnya adalah tahapan penyusunan laporan agar Pemerintah Kabupaten Purwakarta dapat menentukan keputusan yang tepat dalam pengentasan kemiskinan dan penyerapan tenaga kerja melalui UMKM. Hasil dari *clustering* tersebut kemudian akan dilakukan visualisasi berupa pemetaan [25] menggunakan bahasa pemrograman python agar lebih mudah memahami dalam pemetaan ini.



Gambar 6. Pemetaan hasil clustering

## V. KESIMPULAN

Penerapan algoritma k-means dapat mengetahui pengelompokan UMKM di Kabupaten Purwokerto dalam upaya pengentasan kemiskinan dan penyerapan tenaga kerja tahun 2021 berdasarkan lokasi, jumlah UMKM, jumlah penduduk miskin, dan jumlah pencari kerja. Pada pengelompokan didapatkan sebanyak *cluster* prioritas tinggi yaitu sebanyak 6 kecamatan yaitu babakancikao, bungursari, campaka, cibatu, jatiluhur, purwokarta, *cluster* prioritas sedang sebanyak 8 kecamatan yaitu kiarapedes, maniis, pasawahan, pondok salam, sukasari, sukatani, tegal waru, dan wanayasa dan *cluster* prioritas rendah sebanyak 3 kecamatan yaitu; bojong, darangdan, plered. Hasil evaluasi performa pengelompokan UMKM di Kabupaten Purwokerto dalam upaya pengentasan kemiskinan dan penyerapan tenaga kerja tahun 2021 menggunakan algoritma k-means menggunakan *silhouette coefficient* pada penelitian ini pada 3 *cluster* yaitu dengan nilai index 0.45. Visualisasi pemetaan menggunakan library matplotlib.pyplot dan geopandas, bahasa pemrograman python didapatkan hasil pemetaan untuk memudahkan dalam memahami hasil pengelompokan algoritma k-means.

## DAFTAR PUSTAKA

- [1] Fitriani, F.F. Setahun Covid-19 di Indonesia: 2,76 Juta Orang Jadi Warga Miskin. *Bisnis.com* akses 2021, 02 Maret.
- [2] X. Che, B. Zhu, and P. Wang, 'Assessing global energy poverty: An integrated approach', *Energy Policy*, vol. 149, no. C, 2021, Accessed: Aug. 30, 2022. [Online]. Available: <https://ideas.repec.org/a/eee/enepol/v149y2021ics0301421520308107.html>.
- [3] M. Putri, Warga Miskin Bertambah Jadi 74 Juta Orang per September 2021. *Tirto.id*. Akses 16 September 2021
- [4] L.J. Sembiring, RI Dilanda Kemiskinan Ekstrim, Terbanyak Jabar & Jatim! *CNBC Indonesia*. Akses 01 September 2021
- [5] D. H. Jayani, Iklan Lowongan Kerja Menurun pada Agustus 2021. *databoks*.
- [6] I. Wibowo, Efek Pandemi, Pencari Kerja Makin Banyak. *Medcom.id*.
- [7] I. N. Jelita, Pelamar Kerja di Jobstreet Naik 37% selama Pandemi. *Media Indonesia*.
- [8] V. F. Thomas, Susahnya Mencari Kerja saat Pandemi COVID-19 & Pengangguran Naik. *Tirto.id*.
- [9] A. Rahma, Gara-Gara Covid-19, Serapan Tenaga Kerja Industri Pengolahan Turun. *Liputan6*.
- [10] N.Azizah, Nasib angkatan kerja baru di pusaran resesi. *Alinea.id*.
- [11] A. E. Pamuji, Pengaruh pendidikan, pertumbuhan ekonomi, penyerapan tenaga kerja, serta kemiskinan di kabupaten kota provinsi jawa timur tahun 2008-2012. *Jurnal Ekonomi dan Bisnis Islam*.
- [12] Ansori, and J. Priyono, "Analisis Pengaruh Penyerapan Tenaga Kerja dan. Upah Minimum Pekerja terhadap Kemiskinan di Provinsi Jawa Timur Tahun 2009-2015", *Jurnal Ekonomi dan Bisnis*, vol.1 no.3, 2018.
- [13] F. Rozi, Y. Yulmardi, and E. Umiyati, "Pengaruh pertumbuhan ekonomi, upah minimum dan jumlah tenaga kerja terhadap kemiskinan di Kota Jambi Tahun 2000 – 2017", *JELS*, vol. 8, no. 1, pp. 12-25, Mar. 2019.
- [14] N. H. Bustam, 'Pengaruh Jumlah Unit, PDB dan Investasi UMKM Terhadap Penyerapan Tenaga Kerja di Indonesia Periode 2009-2013', *Kutubkhanah*, vol. 19, no. 2, Art. no. 2, Jan. 2017, doi: 10.24014/kutubkhanah.v19i2.2555.
- [15] Y. R. Suci, "Perkembangan Umkm (Usaha Mikro Kecil Dan Menengah) Di Indonesia", *cano*, vol. 6, no. 1, pp. 51–58, Mar. 2017.
- [16] H. Y. Heraldi, N. C. Aprilia, and H. Pratiwi, 'Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means', *Indones. J. Appl. Stat.*, vol. 2, no. 2, Art. no. 2, Dec. 2019, doi: 10.13057/ijas.v2i2.34911.
- [17] P. Sari, B. Pramono, and L. O. H. S. Sagala, 'Improve K-Means Terhadap Status Nilai Gizi Pada Balita', *semanTIK*, vol. 3, no. 1, Art. no. 1, Jun. 2017, doi: 10.55679/semanantik.v3i1.2653.
- [18] D. E. Kurniawan and A. Fatulloh, 'Clustering of Social Conditions in Batam, Indonesia Using K-Means Algorithm and Geographic Information System', *Int. J. Earth Sci. Eng. IJEE*, vol. 10, no. 5, pp. 1076–1080, 2017.
- [19] N. I. Febianto and N. Palasara, 'Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018', *J. Sisfokom Sist. Inf. Dan Komput.*, vol. 8, no. 2, Art. no. 2, Aug. 2019, doi: 10.32736/sisfokom.v8i2.653.
- [20] P. Puntoriza and C. Fibriani, 'Analisis Persebaran UMKM Kota Malang Menggunakan Cluster K-means', *JOINS J. Inf. Syst.*, vol. 5, no. 1, Art. no. 1, May 2020, doi: 10.33633/joins.v5i1.3469.
- [21] D. Astuti, "Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering", *INISTA*, vol. 1, no. 2, pp. 60-72, May 2019.
- [22] I. Sutoyo, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Data Peserta Didik", *pilar*, vol. 14, no. 2, pp. 217-224, 1.
- [23] M. R. Ridlo, S. Defiyanti, A. Primajaya, M. R. Ridlo, S. Defiyanti, and A. Primajaya, 'Implementasi Algoritme K-Means Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang', *CITEE 2017*, pp. 426–433, 2017.
- [24] B. Wira, A. E. Budianto, And A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang", *Rainstek*, Vol. 1, No. 3, Pp. 53 - 68, Sep. 2019.
- [25] A. Ariyanto, D. E. Kurniawan, and A. Fatulloh, 'Rancang Bangun Aplikasi WebGIS untuk Pemetaan Kondisi Sosial Ekonomi Kota Batam', *J. Appl. Inform. Comput. JAIC*, vol. 2, no. 1, Art. no. 1, 2018, doi: 10.30871/jaic.v2i1.904.