

ANALISIS PEMETAAN LEGALITAS USAHA DI KABUPATEN PAMEKASAN MENGGUNAKAN ALGORITMA K-MEANS

Nurul Qamariyah¹, Moh.Aminollah Hamzah², Bakir³

¹Mahasiswa Sistem Informasi, Fakultas Teknik, Universitas Islam Madura (UIM)

^{2,3}Dosen Teknik Informatika, Fakultas Teknik, Universitas Islam Madura (UIM)

¹nurulqamariyah207@gmail.com, ²hamzahmi@gmail.com, ³bakir.madura@gmail.com

ABSTRAK

Pada tahun 2021 Kabupaten Pamekasan menempati pertumbuhan ekonomi tertinggi di Pulau Madura. Dukungan pemerintah melalui legalitas usaha khususnya Nomor Induk Berusaha (NIB) dapat meningkatkan daya saing dan pengembangan Usaha Mikro Kecil dan Menengah (UMKM). Penelitian ini bertujuan penerapan data mining untuk analisis kondisi legalitas UMKM di Kabupaten Pamekasan menggunakan data perizinan NIB sesuai kebijakan terbaru OSS RBA yang terdapat di DPMPTSP Pamekasan. Algoritma k-means dengan teknik clustering dan metodologi CRISP-DM sebagai tahapan standart dalam data mining menghasilkan dua kluster. Kluster 1 yaitu kecamatan dengan jumlah perizinan NIB tertinggi atau UMKM maksimal dan kluster 2 yaitu kecamatan dengan jumlah perizinan NIB terendah atau UMKM minimal. Kluster ini dievaluasi menggunakan Silhouette Coefficient yang menunjukkan penempatan kluster baik untuk data UMK yang bernilai 0,64 (kategori medium structure) dan kluster sangat baik untuk data NonUMK dengan nilai 0,73 (kategori strong structure). Selanjutnya hasil clustering divisualisasikan menggunakan aplikasi QGIS untuk memberikan gambaran terkait legalitas NIB di Kabupaten Pamekasan dan pemerintah terkait dapat merumuskan kebijakan selanjutnya yang lebih tepat untuk mendukung pertumbuhan UMKM.

Kata kunci: UMKM, Legalitas, CRISP-DM, QGIS

ABSTRACT

In 2021, Pamekasan district recorded the highest economic growth on Madura Island. Government support through business legality, especially the Business Identification Number (NIB), can enhance competitiveness and the development of Micro, Small, and Medium Enterprises (MSMEs). This research aims to apply data mining for analyzing the legal conditions of MSMEs in Pamekasan district using NIB licensing data in accordance with the latest OSS RBA policy available at the DPMPTSP Pamekasan. The k-means algorithm with clustering technique and the CRISP-DM methodology as a standard stage in data mining produces two clusters. Cluster 1 refers to sub-districts with the highest number of NIB licenses or maximum MSMEs, and Cluster 2 refers to sub-districts with the lowest number of NIB licenses or minimal MSMEs. This cluster is evaluated using the Silhouette Coefficient which indicates good cluster placement for MSME data with a value of 0.64 (medium structure category) and very good cluster for Non-MSME data with a value of 0.73 (strong structure category). Furthermore, the clustering results are visualized using the QGIS application to provide an overview related to the legality of NIB in Pamekasan district, and the relevant government can formulate more appropriate subsequent policies to support the growth of MSMEs.

Keywords: MSMEs, Legality, CRISP-DM, QGIS

1. PENDAHULUAN

Pada 2021, laju pertumbuhan ekonomi di Kabupaten Pamekasan mengalami pertumbuhan sebesar 3,41 persen yang menjadikan Pamekasan menempati urutan

pertama pertumbuhan ekonomi tertinggi di Pulau Madura. Bupati Pamekasan terus mendorong pertumbuhan ekonomi melalui sektor Usaha Mikro Kecil dan Menengah (UMKM), khususnya dalam membangun

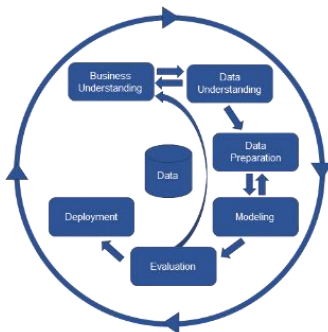
kekamruran sektor ekonomi masyarakat di wilayah setempat (beritajatim.com). Melalui legalitas perizinan usaha, pemerintah dapat memberikan dukungan dan pengembangan usaha yang lebih baik untuk mendorong pertumbuhan usaha. Karena legalitas perizinan usaha ini merupakan salah satu faktor fundamental yang terbukti mempunyai pengaruh signifikan terhadap peningkatan penjualan produk [1]. Pemerintah juga diharuskan menyampaikan informasi terkait perizinan usaha seperti persyaratan yang harus dipenuhi dan tata cara mengajukan [2]. DPMPSTP (Dinas Penanaman Modal Terpadu Satu Pintu) memiliki salah satu tugas dalam penyelenggaraan pelayanan administrasi perizinan dan non perizinan secara terpadu. Adanya website OSS dapat meningkatkan dan memudahkan pelayanan DPMPSTP untuk pengurusan perizinan berusaha yang dapat dilakukan secara aman, cepat dan real time [3]. OSS (One Single Submission) merupakan website layanan perizinan berusaha yang terintegrasi secara elektronik yang disediakan oleh Kementerian Koordinator Bidang Perekonomian Republik Indonesia [4]. Kebijakan perizinan mengalami transformasi dengan berbasis OSS RBA (Single Submission Risk Based Approach) sehingga memudahkan pelaku usaha dalam menerbitkan izin usahanya [5]. OSS RBA ini merupakan perbaikan atau evaluasi dari OSS 1.1 sehingga sifatnya lebih dari sebelumnya [6]. Dilansir radarmadura.jawapos.com (yang terpublish tanggal 21 Desember 2023) hingga saat ini, terdapat 21.704 pelaku usaha mikro kecil menengah (UMKM) yang memiliki Nomor Induk Berusaha (NIB). Padahal, jumlah total pelaku UMKM di Pamekasan mencapai 49.185. Dari hal itu, masih terdapat 27.481 UMKM yang belum memiliki NIB. Kepala Bidang (Kabid) Perizinan dan Non Perizinan DPMPSTP Pamekasan mengungkap sejauh ini masih banyak pelaku usaha yang belum memahami prosedur sistem perizinan usaha melalui OSS sehingga banyak yang belum mendaftarkan produknya atau mengurus NIB (kabarmadura.id). Padahal, legalitas izin usaha sangat krusial bagi perkembangan usaha di era modern ini. Banyak pelaku usaha mikro kecil yang terkendala dalam memperluas bisnisnya serta kesulitan mengakses sumber pendanaan karena belum terdaftar secara resmi dan

memperoleh izin operasional dari pemerintah [7]. Oleh karena itu, untuk meninjau legalitas UMKM di Kabupaten Pamekasan secara lebih efektif, diperlukan analisis yang mendalam terhadap data perizinan NIB. Terdapat penelitian sebelumnya [8] yaitu mencari solusi yang efektif untuk mengurangi tingkat pengangguran dan dapat mengentaskan kemiskinan di Kabupaten Purwakarta sehingga peneliti ini meninjau berbagai penelitian untuk mencari teknik pengelompokan data yang akhirnya ditemukanlah data mining sebagai metode paling relevan untuk pendekatan pengelompokan penelitian ini. Penelitian lain yaitu dari [9] telah menunjukkan bahwa clustering dapat menjadi alat yang berguna untuk memahami karakteristik mahasiswa dan memprediksi waktu kelulusan. Selanjutnya terdapat penelitian [10] yang memiliki tujuan mengidentifikasi pola pembelajaran siswa dan hasil penelitian ini menggunakan metode evaluasi silhouette score yang dapat memberikan panduan dalam menentukan jumlah kluster yang optimal untuk analisis clustering, serta meningkatkan kualitas hasil clustering dan interpretasi data. Maka dari hal itu, pada analisis penelitian ini menerapkan algoritma K-Means untuk Clustering UMKM berdasarkan legalitas perizinan NIB di Kabupaten Pamekasan. [11] menyatakan clustering adalah proses pengelompokan suatu pola yang belum memiliki label dan dilakukan tanpa supervisi menjadi sebuah kelompok yang memiliki karakteristik tertentu. Algoritma K-Means dipilih karena memiliki keunggulannya dalam memproses dataset yang sangat luas sehingga menjadi metode yang paling banyak digunakan dalam berbagai aplikasi data mining [12]. Dalam menerapkan algoritma K-means diperlukan penentuan jumlah cluster yang tepat untuk mengidentifikasi nilai optimalnya [13]. Karena itu, penelitian ini menggunakan validasi cluster dengan *Silhouette Coefficient*. Nilai silhouette dapat digunakan sebagai metrik untuk menilai kualitas hasil clustering dengan menghitungnya pada setiap cluster maupun secara keseluruhan pada suatu algoritma clustering [11]. Hasil cluster analisis ini akan divisualisasikan dengan peta dengan bantuan software QGIS yang diharapkan dapat memberikan gambaran yang lebih jelas mengenai kondisi perizinan NIB di Kabupaten Pamekasan dan menjadi dasar perencanaan

dalam kebijakan yang lebih tepat terkait pentingnya perizinan guna mendukung pertumbuhan UMKM.

2. METODOLOGI PENELITIAN

Menurut [14] CRISP-DM adalah standarisasi dari proses data mining sebagai pemecahan masalah secara umum dalam suatu bisnis atau penelitian. Oleh karena itu, tahapan yang akan digunakan pada penelitian ini menggunakan metodologi seperti gambar 1 berikut :



Gambar 1. Metodologi CRISP-DM

2.1 Bussines Understanding

2.1.1 *Determine business objectives*, yaitu dilakukan pemahaman tujuan penelitian yaitu mengelompokkan UMKM berizin menggunakan data mining clustering k-means melalui tahap wawancara ke pihak Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMPTSP) Pamekasan.

2.1.2 *Assets situation*, pada tahap ini menganalisis fakta dilapangan apakah sudah dilakukan pemetaan legalitas NIB pada UMKM di Kabupaten Pamekasan.

2.1.3 *Determine data mining goals*, proses ini menentukan metode data mining yang akan digunakan dalam penelitian ini. Metode yang akan digunakan yaitu clustering menggunakan algoritma k-means. Algoritma ini dipilih karena bersifat non-hierarkis yang tidak memerlukan pengawasan dalam proses pengelompokkan data.

2.1.4 *Plan activities*, yaitu perencanaan aktivitas yang dirancang untuk mencapai tujuan penelitian serta

menentukan alat dan bahan apa saja yang akan digunakan.

2.2 Data Understanding

Pada tahap kedua ini dilakukan pengumpulan data yang sesuai dengan permasalahan penelitian kemudian dieksplorasi dan diidentifikasi. Data yang digunakan mengenai perizinan NIB pada 4 Agustus 2021 sampai 31 Desember 2024 di Kabupaten Pamekasan berdasarkan catatan yang ada di DPMPTSP.

2.3 Data Preparation

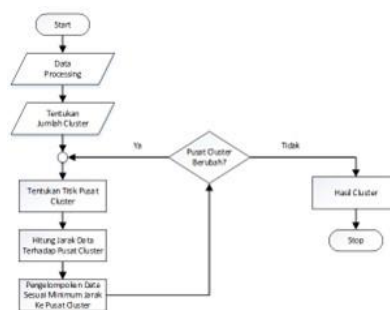
2.3.1 *Data selection* yaitu memilih data dan variabel yang relevan dengan tujuan data mining dalam penelitian. Variabel dalam data yang akan digunakan seperti kecamatan, KBLI, resiko, sektor, dan skala usaha.

2.3.2 *Data preprocessing*, yaitu membersihkan dan mentransformasikan data sehingga sesuai dengan format yang dibutuhkan oleh algoritma. Tahap ini akan dilakukan pengecekan data duplicate dan missing values.

2.3.3 *Transformation*, yaitu tahap pengelompokan atribut atau variabel dalam dataset baru yang sesuai dengan tujuan untuk kemudian diproses dalam data mining. Dataset baru akan dibuat dengan berdasar atribut kecamatan sebagai acuan utama. Pada variabel KBLI akan di lakukan transformation dengan 3 digit yang sudah cukup baik untuk pemodelan yang dibutuhkan.

2.4 Modelling

Pada tahap modelling, penelitian ini akan menggunakan teknik data mining yaitu clustering dengan algoritma k-means terhadap



Gambar 2. Flowchart Algoritma K-Means

data jenis perizinan NIB UMKM dan juga tools google collab dengan bahasa pemrograman python. Berikut gambar 2 yang menjelaskan Algoritma K-Means secara umum yang dilakukan sebagai berikut [15] :

- Tentukan jumlah Cluster/kelompok
- Inisialisasi k pusat klaster (centroid). Pusat-pusat klaster diberi nilai awal dengan angka-angka random dan ditempatkan secara acak.
- Hitung pusat kelompok (centroid) dari data yang ada di masing-masing kelompok.
- Menempatkan masing-masing data dengan centroid klaster yang terdekat. Kedekatan suatu data ke klaster tertentu ditentukan berdasarkan jarak antara data tersebut dengan pusat klaster. Untuk menghitung jarak antara data dengan pusat klaster (centroid) dapat menggunakan jarak Euclidian Distance [16]. Euclidean distance merupakan penghitungan tingkat kesamaan antar titik data dengan berdasar pada jarak [11]. Perhitungan Euclidean distance sebagai berikut :

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (1)$$

Keterangan :

$$d_{ij} = \text{jarak antara objek } i \text{ dengan } j$$

$$C = \text{banyaknya data cluster } C$$

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (3)$$

Hitung nilai silhouette coefficient :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

Hitung nilai silhouette coefficient :

- Centroid akan berpindah dengan menghitung kembali pusat klaster (centroid) dengan keanggotaan klaster menggunakan yang akan membentuk centroid baru. Centroid yang berpindah (centroid baru) merupakan hasil perhitungan rata-rata dari semua data/objek didalam setiap klasternya.
- Proses iterasi ini diulang terus menerus sampai tidak ada perubahan lagi pada kelompok data [17].

2.5 Evaluation

Dari hasil pemodelan, selanjutnya dilakukan tahap evaluation yaitu dengan menggunakan metode silhouette coefficient. Metode silhouette coefficient dipilih karena dapat menilai seberapa baik suatu objek ditempatkan dalam klaster yang tepat. [18] menyatakan silhouette coefficient juga merupakan salah satu metode integrasi validasi clustering yaitu metode cohesion yang berfungsi untuk mengetahui keterkaitan antar objek dalam klaster dan separation yang berfungsi untuk mengetahui jauhnya jarak yang terpisah antar klaster. Langkah silhouette coefficient seperti berikut [19]:

- Hitung rata-rata jarak antara satu titik dengan semua titik lain yang berada di dalam satu cluster atau disebut a(i)

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (2)$$

A = banyaknya data cluster A

- Hitung rata-rata jarak titik dengan semua titik lain yang berada pada cluster lain, kemudian ambil nilai paling minimum dengan persamaan:

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (3)$$

Hitung nilai silhouette coefficient :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

Hitung nilai silhouette coefficient :

Berikut ini nilai silhouette coefficient s(i) menurut Ross (1987) dalam [18] :

- $0.7 < s(i) \leq 1$, Struktur cluster yang kuat (Strong structure)
- $0.5 < s(i) \leq 0.7$, Struktur cluster yang standard (Medium structure)
- $0.25 < s(i) \leq 0.5$, Struktur cluster yang lemah (Weak structure)
- $s(i) \leq 0.25$, Tidak memiliki struktur (No structure)

Kemudian dari hasil evaluasi silhouette coefficient ini dilakukan penentuan langkah selanjutnya apakah bisa dilanjutkan ke tahap berikutnya atau diulang dari awal karena tidak sesuai dengan tujuan penelitian.

2.6 Deployment

Tahap akhir yaitu deployment dimana hasil clustering akan divisualisasikan berupa peta dengan menggunakan tools QGIS (Quantum GIS). Aplikasi QGIS memiliki fitur-fitur yang cukup lengkap, sehingga dapat digunakan oleh pengguna GIS sesuai keperluannya seperti pembuatan peta tematik dan visualisasi data geospasial [20]. Hasil visualisasi pemetaan ini dapat memudahkan pemahaman legalitas NIB pada UMKM di Kabupaten Pamekasan.

3. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan kelompok (cluster) kecamatan-kecamatan di Kabupaten Pamekasan berdasarkan perizinan NIB pada Usaha Mikro, Kecil, dan Menengah (UMKM). Proses klasterisasi ini menggunakan algoritma k-means, dan hasilnya dievaluasi serta divisualisasikan menggunakan perangkat lunak QGIS.

3.1 Business Understanding

3.1.1 *Determine business objectives*, pada tahap awal ini dilakukan observasi melalui wawancara pihak DPMPTSP Pamekasan yaitu Kabid (kepala bidang) perizinan dan non perizinan yang diarahkan ke operator data untuk mencari informasi ketersediaan data yang ada agar data sesuai dengan tujuan penelitian. Fokus observasi ini terkait perizinan NIB UMKM yang ada di setiap kecamatan di Kabupaten Pamekasan.

3.1.2 *Assets Situation*, setelah dilakukannya wawancara dengan bagian operator data DPMPTSP Pamekasan dihasilkan kesimpulan :

- a. Dalam PP No. 5 Tahun 2021 terdapat perubahan perizinan secara elektronik menggunakan sistem Online Single Submission berbasis Risk Based Approach (OSS RBA) yaitu pelaku usaha menyiapkan dokumen dan mengisi data di sistem tanpa harus datang langsung ke Dinas Penanaman Modal dan Pelayanan Terpadu Satu Pintu (DPMTSP) [5].
- b. DPMPTSP Pamekasan belum ada analisis daerah mana saja yang

memiliki tingkat legalitas perizinan NIB yang terbanyak dan terendah.

- c. Data yang ada di DPMPTSP mengenai perizinan NIB di Kabupaten Pamekasan berbentuk excel dengan data UMK dan data Non-UMK pada data terbaru 4 Agustus 2021 sampai 31 Desember 2024.

- d. Visualisasi
Sumber : DPMPTSP Pamekasan penyampaian informasi mengenai UMKM yang memiliki perizinan NIB di Pamekasan belum ada.

3.1.3 *Determine Data Mining Goals*, dari hasil observasi wawancara di DPMPTSP Pamekasan, maka tujuan data mining yang dapat dilakukan dalam penelitian ini yaitu melakukan pengklasteran daerah menggunakan metode clustering dengan teknik algoritma k-means yang selanjutnya akan divisualisasikan berupa peta dengan bantuan aplikasi QGIS untuk mempermudah memahami UMKM yang memiliki perizinan NIB di Kabupaten Pamekasan.

3.2 Data Understanding

Langkah pertama dalam memahami data yang dimiliki, yaitu mengenal data dan menganalisis apa yang bisa dilakukan pada data-data tersebut. Dataset awal yang didapatkan dari DPMPTSP Pamekasan berupa data perizinan NIB di Kabupaten Pamekasan berbentuk excel dengan data terbaru 4 Agustus 2021 - 31 Desember 2024. Terdapat dataset UMK yang berjumlah 21.569 record dan dataset non-UMK berjumlah 165 record. Dataset awal ini memiliki atribut nama perusahaan, KBLI, Judul KBLI, Resiko, Sektor, Skala Usaha, Lokasi Usaha, dan Kecamatan. Dan beberapa atribut private seperti Alamat, No HP, dan Nomor NIB. Dataset awal seperti pada tabel 1 dan 2 berikut:

Tabel 1. Dataset Awal UMK

Tabel 2. Dataset Awal NonUMK

No	Nama Perusahaan	KBLI	Judul KBLI	Risiko	Sektor	Skala Usaha	Lokasi Usaha	Kecamatan
1	CAH AYA AND IL PRO PER TI	'68 11 1	Real Estat Yang Dimil iki Sendi Atau Dise wa	Re nd ah	Perda ganga n	Me ne ng ah	Ke l/D esa Ge lad ak An yar	Pame kasan
...
1	SET ARA MAK	'68 11 1	Real Estat Yang Dimil iki Sendi Atau Dise wa	Me ne ng ah	Peru maha n dan	Me ne ng ah	Ke l/D esa Ce gu k	Tlana kan
6	MUR		Dimil iki Sendi Atau Dise wa	Re nd ah	Kawa san Perm ukim an			
5	GRO UP							

Sumber : DPMPTSP Pamekasan

3.3 Data Preparation

- 3.3.1 *Data Selection*, dilakukan pemilihan data yang dimana pada dataset awal yang didapatkan berisi atribut secara umum, sedangkan pada penelitian ini yang digunakan adalah atribut kecamatan, KBLI, resiko, sektor, dan skala usaha. Untuk itu akan dilakukan penghapusan manual kolom atribut yang tidak digunakan seperti atribut Nama Perusahaan, Judul KBLI, Lokasi Usaha, Alamat, No HP, dan Nomor NIB.
- 3.3.2 *Data Pre-Processing*, pembuatan dataset baru menjadi data yang siap digunakan pada proses data mining. Pembuatan dataset ini dikelompokkan berdasarkan atribut kecamatan sebagai acuan utama dengan atribut jumlah banyaknya data KBLI 3 digit yang sudah cukup baik untuk pemodelan yang dibutuhkan, data risiko, data sektor, dan data skala usaha. Selanjutnya, dilakukan pengecekan

No	Nama Perusahaan	KBLI	Judul KBLI	Risiko	Sektor	Skala Usaha	Lokasi Usaha	Kecamatan
1	M. ZAI NUR RAH MA N	'47 78 4	Perd agan gan Ecer an Bara ng Kera jinan Dari Kera mik	Re nd ah	Pe rd ag an ga n	Mi kr o	Ke l/ De sa Bu la y	Gal is
...
21569	PR. HS - HAR UN SUY ITN O	'12 01 2	Indu stri Roko gi Putih	Ti ng Pe ri n du str ia n		Ke l/ cil sa Ba ja ng	Pa ko ng	

dan penghapusan missing values (data

```
#cek duplicate data
dr.duplicated().sum()
np.int64(0)
```

kosong) dan duplikasi data. Hasilnya gambar 3 dan 4 berikut :

Gambar 4. Cek data duplicate

- 3.3.3 *Transformation*, meringkas informasi dengan pemodelan dataset yang baru yang mengelompokkan berdasarkan kecamatan pada atribut KBLI 3 digit, Skala Usaha, Risiko, dan Sektor dengan metode konvensional, yaitu akumulasi data dari dataset awal. Hasilnya adalah data dikelompokkan sesuai kebutuhan algoritma dengan bersifat numerik sehingga dapat diproses pada data mining seperti pada tabel 3 dan tabel 4 berikut :

Tabel 3. Transformation Data UMK

tabel 6 sebagai berikut :

Tabel 6. Nilai centroid akhir data Non UMK

Centroid	011	012	013	014	016	011	016	011	016	011	016	011	016
1	0	5,0	3	1	2	0,0	...	2	...	0,0	23,5	3	7
2	0	0,0	0	0	0	0,0	...	0	...	0,0	4,0	0	1,15
2	0	8,1	45	81	72	0	...	18	...	36	181	36	363
0	8,2	45	82	72	73	0	...	18	...	36	818	36	64
2	0	0,1	0	0	0	0,1	...	1	...	0,1	4,66	2	3
0	0,67	0	0	5	67	0	...	0	...	0,67	7	0	0

Selanjutnya proses pengelompokan (clustering) hasil data UMK dan Non UMK menggunakan bahasa pyhton yang mengurutkan data hasil cluster 1 dan 2 dimana ditambahkan lagi tampilan nama kecamatan pada hasil clustering. Hasil clustering sebagai berikut :

Gambar 11. Hasil Clustering data UMK

Pada Gambar 11, hasil pengelompokan (clustering) perizinan data UMK menunjukkan adanya dua kelompok (cluster). Kelompok 1 (C1) mengindikasikan bahwa kecamatan Pademawu dan Pamekasan memiliki jumlah perizinan NIB UMK tertinggi (maksimal). Sementara itu, Kelompok 2 (C2) menunjukkan bahwa kecamatan Batumarmar, Galis, Kadur, Larangan, Pakong, Palangaan, Pasean, Pegantenan, Proppo, Tlanakan, dan Waru memiliki jumlah perizinan NIB UMK terendah (minimal). Pada Gambar 12, hasil pengelompokan (clustering) perizinan NonUMK juga menunjukkan adanya dua kelompok (cluster). Kelompok 1 (C1) mengindikasikan bahwa kecamatan Pademawu dan Pamekasan memiliki jumlah perizinan NIB NonUMK tertinggi (maksimal). Sementara itu, Kelompok

2 (C2) menunjukkan bahwa kecamatan Kadur, Larangan, Pakong, Palangaan, Pasean, dan

Gambar 12. Hasil Clustering data Non UMK
Tlanakan memiliki jumlah perizinan NIB terendah (minimal).

3.5 Evaluation

Tahap evaluasi model diperlukan untuk menilai tingkat kualitas pemodelan yang diterapkan, serta memastikan kesesuaiannya dengan tujuan yang telah ditetapkan pada tahap business understanding sebelumnya.

3.5.1 Evaluation Result

Tahap evaluasi hasil dilakukan dengan mengetahui kedekatan relasi antar objek dan seberapa jauh antar cluster terpisah, sehingga diketahui kualitas dari sebuah cluster. Berikut hasil evaluasi nilai silhouette coefficient $s(i)$ menurut Rousseau (1987) pada [18] diatas :

```

# Evaluasi dengan silhouette score
final_silhouette = silhouette_score(df_cleaned, kmeans.labels_)
print("Silhouette score akhir: (final_silhouette)")

# Evaluasi kualitas clustering berdasarkan Silhouette Score
if final_silhouette >= 0.71:
    print("Strong (Kluster terbaik sudah ditemukan)")
elif 0.51 <= final_silhouette < 0.71:
    print("Medium (Penempatan kluster yang wajar)")
elif 0.26 <= final_silhouette < 0.50:
    print("Weak (Strukturanya lemah, coba metode tambahan)")
else:
    print("No Structure (Tidak ada struktur yang ditemukan)")

Silhouette Score akhir: 0.63679964871067
Medium (Penempatan kluster yang wajar)

final_silhouette = silhouette_score(df_cleaned, kmeans.labels_)
print("Silhouette score akhir: (final_silhouette)")

# Evaluasi kualitas clustering berdasarkan Silhouette Score
if final_silhouette >= 0.71:
    print("Strong (Kluster terbaik sudah ditemukan)")
elif 0.51 <= final_silhouette < 0.71:
    print("Medium (Penempatan kluster yang wajar)")
elif 0.26 <= final_silhouette < 0.50:
    print("Weak (Strukturanya lemah, coba metode tambahan)")
else:
    print("No Structure (Tidak ada struktur yang ditemukan)")

Silhouette Score akhir: 0.72886802457841
Strong (Kluster terbaik sudah ditemukan)
    
```

Gambar 14. Hasil Evaluasi Non UMK

Pada gambar 13 menjelaskan bahwa hasil evaluasi data UMK menggunakan silhouette coefficient atau silhouette score menunjukkan kategori medium dengan nilai $s(i)$ yaitu 0,64.

Pada gambar 14 menjelaskan bahwa hasil evaluasi data NonUMK menggunakan silhouette coefficient atau silhouette score menunjukkan kategori Strong dengan nilai $s(i)$ yaitu 0,73.

Gambar 13. Hasil Evaluasi UMK

3.5.2 Determine Next Step

Tahap ini merupakan penentuan langkah selanjutnya berdasarkan hasil evaluasi. Terdapat dua opsi dalam

- yang tercatat dalam dataset UMK (Usaha Mikro Kecil) dan Non-UMK.
- Hasil pengelompokan daerah perizinan NIB di Kabupaten Pamekasan, dengan kecamatan sebagai acuan utama, menunjukkan pola yang bervariasi antara dataset UMK dan Non-UMK. Pada dataset UMK, dua kecamatan tergolong ke dalam Klaster 1 (maksimal) dengan legalitas NIB terbanyak, sementara sebelas kecamatan berada di Klaster 2 (minimal), mengindikasikan perlunya upaya lebih dalam penggalakan perizinan NIB. Untuk dataset Non-UMK, dua kecamatan masuk Klaster 1 (maksimal) dan enam kecamatan di Klaster 2 (minimal); menariknya, ada lima kecamatan yang berstatus "undefined", menandakan belum adanya perizinan Non-UMK yang tercatat, kemungkinan karena ketiadaan entitas atau belum adanya proses perizinan di wilayah tersebut.
 - Evaluasi pemetaan legalitas Nomor Induk Berusaha (NIB) di Kabupaten Pamekasan, yang menggunakan algoritma K-Means, telah dievaluasi kualitasnya melalui Silhouette Coefficient. Hasilnya menunjukkan bahwa pada data UMK, nilai 0,63 mengindikasikan "medium structure", yang berarti penempatan cluster cukup wajar. Sebaliknya, pada data Non-UMK, nilai 0,72 menunjukkan "strong structure", menegaskan bahwa penempatan cluster sudah sangat kuat dan optimal, bahkan cluster terbaik telah berhasil ditemukan.

UCAPAN TERIMA KASIH

Terima kasih saya sampaikan kepada Universitas Islam Madura termasuk para dosen, Dekan, dan Kaprodi atas ilmu yang telah diberikan sehingga ide penelitian ini dapat terwujud. Terima kasih juga kepada DPMPSTP Pamekasan atas dukungan yang diberikan dalam pelaksanaan penelitian ini. Terima kasih juga kepada Bapak Hamzah selaku dosen pembimbing 1 dan Bapak Bakir selaku dosen pembimbing 2 yang telah membimbing penelitian ini. Tak lupa, dukungan motivasi yang tak ternilai dari orang-orang terdekat, terutama keluarga dan sahabatku yang senantiasa memberikan semangat di kala lelah dan jenuh, juga saya ucapkan terima kasih sebesar-besarnya.

DAFTAR PUSTAKA

- [1] H. Akim, "Pengaruh Legalitas Usaha, Packaging produk dan Labelisasi Halal MUI terhadap Peningkatan Volume Penjualan Produk, Usaha Mikro Kecil Menengah (UMKM) Di Kecamatan Cilacap Selatan Kabupaten Cilacap," *Jurakunman (Jurnal Akunt. dan Manajemen)*, vol. 16, no. 1, p. 118, 2023, doi: 10.48042/jurakunman.v16i1.209.
- [2] A. Redi, L. Marfugah, R. F. Fansuri, M. Prawira, and A. Lafentia, "Perizinan Usaha Mikro, Kecil, Dan Menengah (Umkm): Bentuk Pemberdayaan, Perlindungan Hukum Dan Mewujudkan Negara Kesejahteraan," *J. Muara Ilmu Sos. Humaniora, dan Seni*, vol. 6, no. 1, p. 282, 2022, doi: 10.24912/jmishumsen.v6i1.13553.2022.
- [3] F. P. Rahayu, E. Paselle, and T. R. Khaerani, "Pelaksanaan Pelayanan Perizinan Berusaha Online Single Submission (OSS) Di DPMPSTP Kota Samarinda," *eJournal Adm. Publik*, vol. 9, no. 2, pp. 5018–5032, 2021.
- [4] Nashrudin Latif, Suharyanto, Bayu Adi, Untung Lasiyono, and Mohamad Afrizal Miradji, "Digitalisasi Pengajuan Izin Usaha UMKM di Desa Banjarkemantren, Kecamatan Buduran, Kabupaten Sidoarjo," *Ekobis Abdimas J. Pengabd. Masy.*, vol. 2, no. 1, pp.

- 92–101, 2021, doi: 10.36456/ekobisabdimas.2.1.3911.
- [5] R. Bahroni, M. Rumbekwan, and A. H. Ginting, "IMPLEMENTASI KEBIJAKAN PELAYANAN PERIZINAN USAHA MIKRO KECIL BERBASIS ONLINE SINGLE SUBMISSION RISKA BASED APPROACH (OSS RBA) DI KABUPATEN SITUBONDO PROVINSI JAWA TIMUR Institut Pemerintahan Dalam Negeri IMPLEMENTATION OF POLICY FOR LICENSING OF SMALL M," *J. Ilm. Wahana Bhakti Praja*, vol. 12, no. 1, pp. 71–85, 2022.
- [6] T. Lestaringtyas and M. Roqib, "Perlindungan Data Pribadi Pengguna Sistem Layanan Perizinan Berusaha Terintegrasi Secara Elektronik Oss 1.1 Dan Oss Rba (Risk Basic Approach)," *J. Jendela Huk.*, vol. 8, no. 2, pp. 25–34, 2021, doi: 10.24929/fh.v8i2.1576.
- [7] N. Soimah and D. Q. Imelda, "Urgensi Legalitas Usaha Bagi UMKM," *J. Benuanta*, vol. 2, no. 1, pp. 21–25, 2023, doi: 10.61323/jb.v2i1.47.
- [8] H. Kurniadewi, R. A. Hakim, M. Jajuli, and J. H. Jaman, "Pemetaan UMKM dalam Upaya Pengentasan Kemiskinan dan Penyerapan Tenaga Kerja Menggunakan Algoritma K-Means," *J. Appl. Informatics Comput.*, vol. 6, no. 2, pp. 113–119, 2022, doi: 10.30871/jaic.v6i2.4227.
- [9] A. Khaerunnisa, "Analisis Tingkat Kelulusan Mahasiswa di Unisba dengan menggunakan Algoritma K-Means Clustering," *J. Ris. Mat.*, pp. 67–76, 2022, doi: 10.29313/jrm.v2i1.1018.
- [10] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. Dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024, [Online]. Available: <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- [11] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, p. 48, 2021, doi: 10.51557/pt_jiit.v6i2.659.
- [12] S. Pujiono, R. Astuti, and F. Muhamad Basysyar, "Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 615–620, 2024, doi: 10.36040/jati.v8i1.8360.
- [13] R. S. Wicaksana, D. Heksaputra, T. A. Syah, and F. F. Nur'aini, "Pendekatan K-Means Clustering Metode Elbow Pada Analisis Motivasi Pengunjung Festival Halal JHF#2," *J. Ilm. Ekon. Islam*, vol. 9, no. 3, p. 4162, 2023, doi: 10.29040/jiei.v9i3.10591.
- [14] R. N. Fahmi, M. Jajuli, and N. Sulistiyowati, "Analisis Pemetaan Tingkat Kriminalitas di Kabupaten Karawang menggunakan Algoritma K-Means," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 4, no. 1, pp. 67–79, 2021, doi: 10.31539/intecom.s.v4i1.2413.
- [15] A. Bahauddin, A. Fatmawati, and F. Permata Sari, "Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means," *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.
- [16] A. Septianingsih, "Analisis K-Means Clustering Pada Pemetaan Provinsi Indonesia Berdasarkan Indikator Rumah Layak Huni," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 3, no. 1, pp. 224–241, 2022, doi: 10.46306/lb.v3i1.116.
- [17] F. Salsabila, I. Fitrianti, Y. Umaidah, and N. Heryana, "Penerapan Metode Crisp-Dm Untuk Analisa Pendapatan Bersih Bulanan Pekerja Informal Di Provinsi Jawa Barat Dengan Algoritma K-Means," *Dinamik*, vol. 28, no. 2, pp. 97–104, 2023, doi: 10.35315/dinamik.v28i2.9454.
- [18] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli,

-
- “Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM,” *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [19] Y. Kustiyahningsih, A. Khozaimi, and J. Purnama, “Pengelompokan UMKM Batik Madura Menggunakan Metode K-Means dan Silhouette Coefficient,” *Teknika*, vol. 13, no. 2, pp. 192–198, 2024, doi: 10.34148/teknika.v13i2.779.
- [20] Aliya Raihana, Zaky Ahmad Faisal, Muhammad Fiqih, Irna Aulia, Muhamad Yazid Imani, and Walidatush Sholihah, “Pemanfaatan Aplikasi Qgis Untuk Pemetaan Toko Obat Hama,” *JSiI (Jurnal Sist. Informasi)*, vol. 11, no. 1, pp. 10–14, 2024, doi: 10.30656/jsii.v11i1.7901.