

PENERAPAN ALGORITMA K-PROTOTYPES DALAM ANALISIS POTENSI UMKM DI KABUPATEN ASAHAN

William Ramdhan¹, Nurwati², Santoso³, Dava Erlangga⁴

Universitas Royal, Kisaran

email: ¹william.ramdhan052@gmail.com, ²nurwati763@gmail.com,

³massantoso78@gmail.com, ⁴davaerlangga37@gmail.com

Abstract: *Micro, Small, and Medium Enterprises (MSMEs) in Asahan Regency have very heterogeneous characteristics, both in terms of business sectors, capital size, and income achievement. This study aims to cluster MSMEs using the K-Prototypes algorithm to group business actors into uniform clusters based on numerical and categorical characteristics. The methodology used includes the pre-processing stage, variable transformation, determining the optimal number of clusters using the elbow method, and implementing the K-Prototypes algorithm. The results of the study showed that five main clusters were successfully formed, each showing a different pattern in terms of capital, net income, and dominant business sector. Data visualization and exploration (EDA) also strengthened the understanding of the cluster structure that was formed. The cluster with the highest capital and income is dominated by the medium-scale trade sector, while the cluster with the lowest capital and income is identical to micro MSMEs in the culinary and service sectors. These findings prove that the K-Prototypes algorithm is effective in identifying MSME segmentation in a more structured manner and can be the basis for designing more targeted MSME development strategies.*

Keyword: *UMKM; clustering; K-Prototypes; mixed data; segmentation analysis.*

Abstrak: Usaha Mikro, Kecil, dan Menengah (UMKM) di Kabupaten Asahan memiliki karakteristik yang sangat heterogen, baik dari sisi sektor usaha, besaran modal, hingga capaian income. Penelitian ini bertujuan untuk melakukan klasterisasi UMKM menggunakan algoritma K-Prototypes guna mengelompokkan pelaku usaha ke dalam klaster-klaster yang seragam berdasarkan karakteristik numerik dan kategorikal. Metodologi yang digunakan mencakup tahap pre-processing, transformasi variabel, penentuan jumlah klaster optimal menggunakan metode elbow, serta implementasi algoritma K-Prototypes. Hasil penelitian menunjukkan bahwa lima klaster utama berhasil dibentuk, masing-masing menunjukkan pola yang berbeda dalam hal modal, income bersih, dan sektor usaha dominan. Visualisasi dan eksplorasi data (EDA) turut memperkuat pemahaman terhadap struktur klaster yang terbentuk. Klaster dengan modal dan income tertinggi didominasi oleh sektor perdagangan skala menengah, sedangkan klaster dengan modal dan income terendah identik dengan UMKM mikro di sektor kuliner dan jasa. Temuan ini membuktikan bahwa algoritma K-Prototypes efektif digunakan untuk mengidentifikasi segmentasi UMKM secara lebih terstruktur dan dapat menjadi dasar dalam merancang strategi pengembangan UMKM yang lebih tepat sasaran.

Kata kunci: UMKM; klasterisasi; K-Prototypes; data campuran; analisis segmentasi

PENDAHULUAN

Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan tulang punggung perekonomian nasional, baik di negara maju maupun berkembang (Aliyah, 2022).

Di Indonesia, UMKM menyumbang sekitar 60% terhadap Produk Domestik Bruto (PDB) dan lebih dari 97% terhadap total lapangan kerja. Peran UMKM tidak hanya terbatas pada aspek ekonomi, tetapi juga sosial, khususnya dalam hal

pemerataan pendapatan, pengurangan pengangguran, serta penguatan struktur ekonomi berbasis kerakyatan (Wita Setiawati et al., 2024). Di tingkat daerah, seperti Kabupaten Asahan di Provinsi Sumatera Utara, UMKM berkembang secara signifikan dan menjadi motor penggerak utama ekonomi lokal.

Namun demikian, dinamika pertumbuhan UMKM di daerah tidak lepas dari berbagai tantangan, terutama dalam hal keterbatasan data yang terintegrasi, kurangnya segmentasi yang jelas, serta minimnya intervensi kebijakan yang berbasis pada analisis mendalam (Asyifah Fauzah Delfira et al., 2025). UMKM di Kabupaten Asahan memiliki karakteristik yang sangat heterogen, baik dari sisi jenis usaha, modal maupun skala pendapatan. Heterogenitas ini menyulitkan proses perumusan kebijakan yang bersifat menyeluruh dan tepat sasaran, karena pendekatan yang bersifat umum tidak mampu menjawab kebutuhan spesifik dari tiap segmen UMKM. Oleh karena itu, diperlukan metode analisis yang mampu mengelompokkan UMKM berdasarkan karakteristik dominan yang dimiliki secara objektif dan sistematis.

Pada analisis data, proses klusterisasi menjadi salah satu pendekatan yang relevan untuk menjawab permasalahan tersebut. Keberadaan data UMKM bersifat campuran yang terdiri dari atribut numerik (modal dan omset) dan kategorikal (seperti jenis usaha) untuk itu perlunya pendekatan yang mampu menangani data tersebut yakni algoritma K-Prototypes merupakan solusi dalam menangani data numerik dan kategorikal secara bersamaan (Purba, Ariani Fitri; Mustafid & Kartikasari, 2025). K-Prototypes merupakan penggabungan dari algoritma K-Means dan K-Modes, sehingga sangat sesuai untuk klusterisasi data UMKM yang kompleks dan variatif (Saniyah et al., 2024).

Urgensi dari riset ini terletak pada kebutuhan mendesak untuk menyediakan landasan empiris yang kuat dalam menyusun strategi pengembangan

UMKM berbasis data. Dengan penerapan algoritma K-Prototypes, proses klusterisasi diharapkan dapat menghasilkan segmentasi UMKM yang lebih akurat dan representatif, yang pada gilirannya dapat dimanfaatkan oleh pemerintah daerah, lembaga keuangan, maupun pelaku industri dalam merancang kebijakan dan program pemberdayaan UMKM yang tepat guna. Selain itu, penelitian ini juga berkontribusi dalam pengembangan pendekatan data mining berbasis algoritma pada konteks kebijakan publik daerah, khususnya dalam pemetaan potensi ekonomi mikro yang belum banyak tersentuh analisis ilmiah mendalam.

METODE

K-Prototypes merupakan algoritma klusterisasi yang dikembangkan untuk menangani data dengan atribut campuran, yaitu data numerik dan kategorikal. Algoritma ini merupakan perluasan dari algoritma K-Means (yang hanya cocok untuk data numerik) dan K-Modes (yang digunakan untuk data kategorikal). K-Prototypes memadukan keunggulan keduanya dengan memodifikasi fungsi jarak agar dapat mengukur perbedaan antara data numerik dan kategorikal secara bersamaan (Novidianto & Fithriasari, 2021).

K-Prototypes bertujuan untuk mempartisi data ke dalam sejumlah k kluster sedemikian rupa sehingga jarak total (total dissimilarity) antara objek dalam kluster dengan prototipe (centroid) klasternya diminimalkan (Reihanah et al., 2024). Untuk data campuran, jarak $D(x,y)$ antara dua objek didefinisikan sebagai:

$$D_{(x,y)} = \pi \sum_{j \in \text{numerik}} (x_j - y_j) + \gamma \sum_{j \in \text{kategorikal}} \delta(x_j, y_j)$$

Dimana

x_j dan y_j : nilai fitur ke $-j$

$\delta(x_j, y_j) = 0$ jika $x_j = y_j$, dan 1 jika $x_j \neq y_j$

γ : parameter bobot untuk menyeimbangkan kontribusi atribut kategorikal terhadap fungsi jarak.

Tahapan umum dalam algoritma K-Prototypes:

1. Inisialisasi
Menentukan jumlah kluster k dan memilih prototipe awal secara acak dari data. Setiap prototipe terdiri atas nilai rata-rata (untuk atribut numerik) dan modus (untuk atribut kategorikal).
2. Penghitungan Jarak dan Pengelompokan
Menghitung jarak setiap objek terhadap semua prototipe menggunakan fungsi dissimilarity gabungan. Objek kemudian ditempatkan ke dalam kluster dengan prototipe terdekat.
3. Pembaruan Prototipe untuk setiap kluster
Atribut numerik diperbarui dengan menghitung rata-rata dari anggota kluster.
Atribut kategorikal diperbarui dengan mencari modus (nilai yang paling sering muncul).
4. Iterasi
Langkah pengelompokan dan pembaruan prototipe diulang hingga tidak terjadi perubahan signifikan dalam keanggotaan kluster atau jumlah iterasi maksimum tercapai.

5. Evaluasi
Setelah proses konvergen, hasil klusterisasi dievaluasi menggunakan metrik internal seperti *cost function*, *silhouette score*, atau *within-cluster dissimilarity*.

HASIL DAN PEMBAHASAN

Proses KDD diawali dengan proses data selection/pengumpulan data menggunakan data primer yang tersedia pada Dinas koperasi perdagangan dan Perindustrian Kab. Asahan. Ketersediaan data pada dinas tersebut meliputi nama pemilik, jenis UMKM, Modal, Income/Pendapatan.

Pre-processing

Pada tahap pre-processing, dilakukan identifikasi terhadap kemungkinan adanya nilai yang hilang maupun data ganda. Berdasarkan hasil analisis pre-processing, dapat dipastikan bahwa seluruh data dalam kondisi lengkap dan tidak ditemukan adanya duplikasi. Dinas koperasi perdagangan dan perindustrian Kab. Asahan setelah proses pre-processing

Tabel 1 Dataset UMKM Kab. Asahan

No	Nama Pemilik	Jenis UMKM	Modal	Income
1	Rosma Sibarani	Kedai Kelontong	6000000	3000000
2	Tommy Paisal	Steam Mobil Dan Motor	12500000	5000000
3	Syafrida Yani	Fashion	10000000	5000000
4	Ripka Pransiska Br Rajagukguk	Pertanian	6000000	2000000
5	Evi Palentina Rajagukguk	Pertanian	12000000	4500000
7	Nurul Aulia	Pertanian	25000000	10000000
...				
308	Andi Wahyudi	Jasa Layanan Antar	10000000	7000000
309	Turiman	Pertanian	15000000	8500000

Transformasi Data

Ketersediaan data yang heterogenitas pada dataset UMKM maka

diperlukan variabel operasional untuk memastikan kejelasan dan konsistensi dalam proses pengumpulan data. Tanpa

definisi yang jelas, terdapat risiko kesalahan akibat penggunaan instrumen yang kurang tepat atau tidak konsisten. Oleh karena itu, setiap variabel yang akan

diukur perlu dirumuskan secara terperinci melalui batasan-batasan yang spesifik yakni data jenis UMKM seperti yang terlihat pada Tabel 2.

Tabel 2 Operasional Variabel

No	Sektor UMKM	Jenis UMKM
1	Perdagangan	Usaha yang bergerak dalam penjualan barang, baik baru maupun bekas, seperti toko kelontong, pedagang kaki lima, dan usaha perdagangan di pasar.
2	Pertanian	Meliputi berbagai subsektor seperti tanaman pangan, hortikultura (sayuran, buah-buahan, tanaman hias), perkebunan, perikanan, dan peternakan
3	Jasa	Berbagai jenis layanan seperti servis kendaraan, servis elektronik, salon kecantikan, laundry, dan jasa fotografer
4	Kuliner	Usaha mencakup berbagai jenis makanan dan minuman tradisional maupun modern.

Transformasi nilai variabel "Jenis UMKM" menjadi nilai "Sektor UMKM" pada Tabel 3 dilakukan dengan cara mengkodekan masing-masing kategori yang memungkinkan pemrosesan oleh algoritma K-Prototypes, sesuai prinsip kerja K-Prototypes yang membedakan perlakuan antara atribut numerik dan kategorikal.

Tabel 3 Dataset Transformasi UMKM Kab. Asahan

No	Sektor UMKM	Modal	Income
1	Perdagangan	6000000	3000000
2	Jasa	12500000	5000000
3	Perdagangan	10000000	5000000
4	Pertanian	6000000	2000000
5	Pertanian	12000000	4500000
7	Pertanian	25000000	10000000
...			
308	Jasa	10000000	7000000
309	Pertanian	15000000	8500000

Data Mining

Data mining merupakan proses untuk menggali pola atau pengetahuan tersembunyi dari sekumpulan data dan mengubahnya menjadi informasi yang

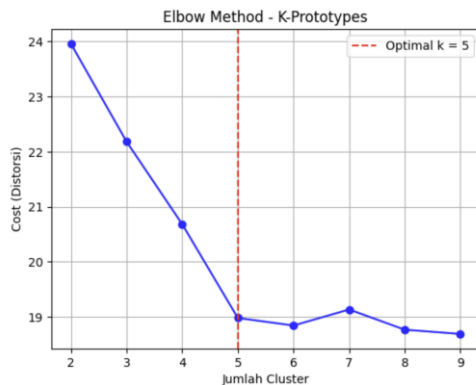
bermakna (Apriyani et al., 2023). Setelah data melalui tahap *preprocessing* dan transformasi, langkah selanjutnya adalah mengidentifikasi pola dan membentuk kluster berdasarkan karakteristik tertentu.

Salah satu algoritma yang digunakan dalam teknik *clustering* adalah K-Prototypes, yaitu metode pengelompokan non-hierarki yang dirancang untuk menangani data campuran, baik numerik maupun kategorikal. Algoritma ini mensyaratkan penentuan jumlah kluster (*k*) di awal proses, dan pengelompokan dilakukan berdasarkan kesamaan terhadap variabel penciri, sehingga objek-objek dalam satu kluster memiliki karakteristik yang serupa (Apriyani et al., 2023).

Penentuan jumlah kluster menjadi aspek yang sangat penting karena jumlah yang berbeda dapat menghasilkan segmentasi dan interpretasi yang berbeda pula (Kustanto et al., 2024). Oleh karena itu, untuk memperoleh kluster yang optimal dan benar-benar merepresentasikan struktur alami dalam data, digunakan pendekatan elbow method, yaitu dengan melihat titik tekuk pada grafik evaluasi sebagai indikator jumlah kluster terbaik (Rahmawati et al., 2024).

```
costs = []
K = range(2, 10)
for k in K:
    kproto = KPrototypes(n_clusters=k, init='Cao', verbose=0, random_state=42)
    kproto.fit_predict(X_matrix, categorical_cat_idx)
    costs.append(kproto.cost_)
    k1 = KneedleLocator(K, costs, curves='convex', direction='decreasing')
    optimal_k = k1.elbow

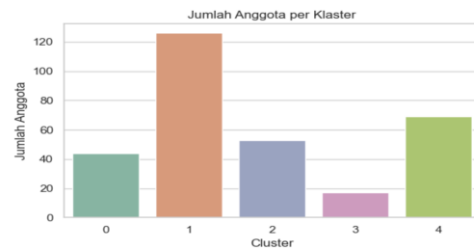
# Plot (optional)
plt.plot(K, costs, 'bo-')
plt.xlabel('Jumlah Cluster')
plt.ylabel('Cost (Distorsi)')
plt.title('Elbow Method - K-Prototypes')
plt.axvline(x=optimal_k, color='red', linestyle='--', label=f'Optimal k = {optimal_k}')
plt.legend()
plt.grid(True)
plt.show()
```



Gambar 1 Penentuan Jumlah Kluster dengan pendekatan Elbow Method

Proses selanjutnya melakukan klusterisasi menggunakan metode K-Prototypes dengan jumlah kluster optimal yang ditentukan melalui pendekatan elbow yakni $K=5$. Hasil mining tersebut menghasilkan lima kluster utama. Setiap kluster menunjukkan karakteristik unik

berdasarkan kombinasi variabel numerik (modal dan income) serta variabel kategorikal (sektor usaha). Jumlah UMKM pada masing-masing kluster menunjukkan distribusi yang cukup berimbang, dengan Kluster 1 sebagai kelompok terbesar (126 UMKM) dan Kluster 3 sebagai yang terkecil (17 UMKM).



Gambar 2 Visualisasi Jumlah Anggota Per Kluster

Tabel 4 Hasil Cluster Optimum Berdasarkan Elbow Methods

Nama Usaha	Sektor Usaha	Modal	Income	Cluster
A1	Perdagangan	6000000	3000000	0
A2	Jasa	12500000	5000000	1
A3	Perdagangan	10000000	5000000	1
A4	Pertanian	6000000	2000000	2
...				
A308	Jasa	10000000	7000000	4
A309	Pertanian	15000000	8500000	2

Exploratory Data Analysis (EDA)

Proses Exploratory Data Analysis (EDA) berperan penting dalam memastikan bahwa data UMKM yang digunakan untuk klusterisasi mencerminkan kondisi nyata dan layak dianalisis. Mengingat karakteristik UMKM yang sangat heterogen, EDA membantu memahami distribusi modal dan income, mendeteksi outlier, serta melihat kecenderungan umum antar pelaku usaha. Melalui visualisasi dan analisis statistik, EDA memungkinkan identifikasi variabel penting, transformasi data, dan pemahaman pola sektoral sebelum pembentukan kluster dilakukan (Ramdhan et al., 2025).

Selain itu, hasil EDA juga memperkuat interpretasi kluster akhir,

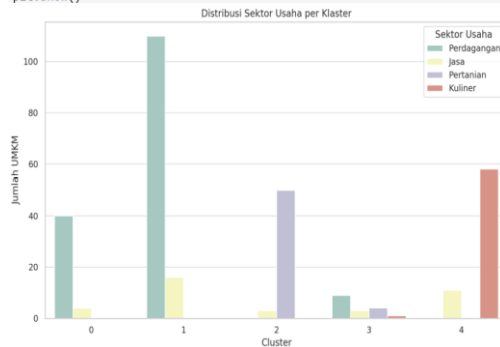
sehingga kluster yang terbentuk dapat dikaitkan dengan kondisi nyata di lapangan. Dengan demikian, EDA menjadi fondasi penting dalam menghasilkan analisis yang relevan dan mendukung pengambilan keputusan berbasis data dalam pengembangan UMKM.

Visualisasi hasil *Exploratory Data Analysis* (EDA) berperan penting dalam memahami struktur dan karakteristik tiap kluster UMKM. Tiga visualisasi utama meliputi distribusi sektor usaha, boxplot modal, dan boxplot income. Visualisasi ini membantu mengidentifikasi untuk mengungkapkan hubungan antara kondisi ekonomi pelaku UMKM dan sektor usaha dominan dalam tiap kluster. Dengan

demikian, hasil klasterisasi tidak hanya bersifat statistik, tetapi juga memiliki

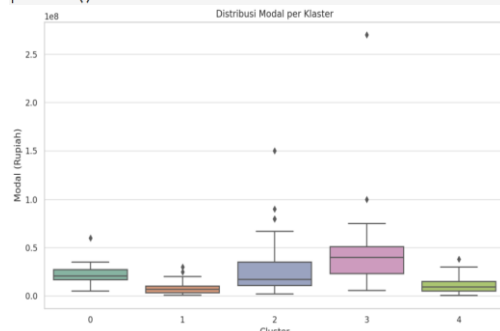
makna sektoral yang strategis.

```
plt.figure(figsize=(10, 6))
sns.countplot(data=df, x='Cluster', hue='Sektor_Usaha', palette='Set3')
plt.title('Distribusi Sektor Usaha per Kluster')
plt.xlabel('Cluster')
plt.ylabel('Jumlah UMKM')
plt.legend(title='Sektor Usaha')
plt.tight_layout()
plt.show()
```



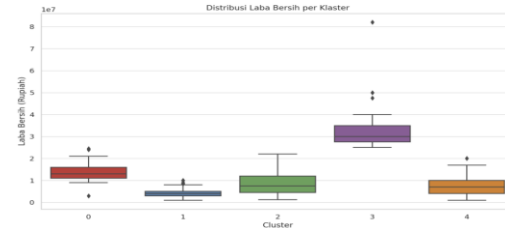
Gambar 3 Distribusi Sektor Usaha per Kluster

```
plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='Cluster', y='Modal', palette='Set2')
plt.title('Distribusi Modal per Kluster')
plt.xlabel('Cluster')
plt.ylabel('Modal (Rupiah)')
plt.tight_layout()
plt.show()
```



Gambar 4 Boxplot Modal per Kluster

```
plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='Cluster', y='Laba_Bersih', palette='Set1')
plt.title('Distribusi Laba Bersih per Kluster')
plt.xlabel('Cluster')
plt.ylabel('Laba Bersih (Rupiah)')
plt.tight_layout()
plt.show()
```



Gambar 5. Boxplot Income per Kluster

Hasil eksplorasi data terhadap kluster yang terbentuk menunjukkan adanya pola distribusi yang menarik, baik dari sisi sektor usaha, modal, maupun income. Berdasarkan distribusi sektor usaha per kluster, terlihat bahwa Klaster 1 dan Klaster 4 didominasi oleh pelaku UMKM di sektor perdagangan dan jasa, yang secara umum merepresentasikan kelompok usaha berskala mikro dan kecil. Di sisi lain, Klaster 0 dan Klaster 2 menunjukkan komposisi sektor yang lebih beragam, termasuk keberadaan sektor pertanian yang cukup signifikan. Sementara itu, Klaster 3 yang jumlah anggotanya paling sedikit justru cenderung terdiri atas sektor-sektor usaha yang padat modal, mencerminkan jenis UMKM yang memiliki kapasitas finansial lebih besar.

Dari sisi struktur permodalan, hasil boxplot menunjukkan bahwa Klaster 3 merupakan kluster dengan nilai modal tertinggi, bahkan mencakup UMKM dengan modal hingga ratusan juta rupiah. Sebaliknya, Klaster 1 dan Klaster 4 lebih dominan pada kelompok usaha dengan modal kecil hingga menengah, umumnya di bawah Rp15 juta. Klaster 0 dan Klaster 2 menunjukkan tingkat variasi modal yang cukup luas, dengan kecenderungan berada di kelompok menengah.

Analisis terhadap income juga menunjukkan pola yang selaras dengan distribusi modal. Klaster 3, yang memiliki nilai modal tinggi, juga menunjukkan income yang sangat besar, mencerminkan UMKM yang berada pada fase pertumbuhan atau bahkan mendekati skala menengah. Klaster 1 menempati posisi terendah dalam hal income, dengan sebaran yang relatif merata, sesuai dengan karakteristik usaha mikro. Adapun Klaster 0, Klaster 2, dan Klaster 4 memiliki distribusi income yang berada pada

kisaran menengah ke atas, dengan beberapa outlier yang menunjukkan potensi usaha lebih besar.

Temuan ini mengindikasikan bahwa hasil klasterisasi tidak hanya berhasil mengelompokkan UMKM berdasarkan nilai numerik seperti modal dan income, tetapi juga mencerminkan keterkaitan erat antara sektor usaha dan kapasitas ekonomi masing-masing pelaku. Dengan demikian, hasil ini dapat dijadikan dasar untuk menyusun strategi intervensi yang lebih terarah sesuai dengan karakteristik klaster masing-masing

SIMPULAN

Hasil klasterisasi UMKM Kabupaten Asahan menggunakan algoritma K-Prototypes menunjukkan efektivitas metode ini dalam mengelompokkan data campuran numerik dan kategorikal secara optimal. Melalui pendekatan elbow method, ditentukan bahwa jumlah klaster terbaik adalah lima, dengan masing-masing klaster merepresentasikan struktur usaha yang berbeda secara signifikan.

Klaster 1 dan Klaster 4 didominasi oleh sektor perdagangan dan kuliner dengan karakteristik modal dan income

yang rendah, mengindikasikan kelompok usaha skala mikro. Klaster 2 memiliki rata-rata modal dan income menengah dengan dominasi sektor pertanian, yang menunjukkan potensi pengembangan pada bidang agribisnis. Sementara itu, Klaster 3 menunjukkan karakteristik paling kuat dari sisi finansial, dengan nilai modal dan income tertinggi, didominasi oleh usaha perdagangan dalam skala lebih besar yang cenderung berada pada fase pertumbuhan dan ekspansi.

Distribusi sektor usaha dalam masing-masing klaster tidak hanya memperjelas segmentasi struktural UMKM di lapangan, tetapi juga memberikan kontribusi praktis dalam penyusunan strategi pengembangan berbasis data. Pemahaman terhadap sektor dominan dalam tiap klaster memungkinkan perumusan intervensi yang lebih terarah, seperti dukungan aspek produksi dan distribusi pada klaster sektor pertanian, atau penguatan kualitas produk dan kemasan pada klaster kuliner.

Oleh karena itu, integrasi antara algoritma K-Prototypes dan eksplorasi distribusi sektor usaha memberikan kerangka kerja yang komprehensif untuk mendukung pengambilan kebijakan strategis dalam pemberdayaan UMKM daerah secara objektif, adaptif, dan berbasis bukti.

DAFTAR PUSTAKA

- Aliyah, A. H. (2022). Peran Usaha Mikro Kecil dan Menengah (UMKM) untuk Meningkatkan Kesejahteraan Masyarakat. *WELFARE Jurnal Ilmu Ekonomi*, 3(1), 64–72. <https://doi.org/10.37058/wlfr.v3i1.4719>
- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20–33. <https://doi.org/10.56211/helloworld.v2i1.230>
- Asyifah Fauzah Delfira, Putri Yuliarman, Muhammad Baihaqi, & Muhammad Baihaqi. (2025). Analisis Permasalahan Perkembangan UMKM Terhadap Perekonomian di Indonesia. *Jurnal Ekonomi, Bisnis Dan Manajemen*, 4(1), 167–180. <https://doi.org/10.58192/ebismen.v4i1.3184>
- Kustanto, Y., Arumi, R., Sasongko, D., Ully Artha, E., & Prabowo, N. A. (2024). Implementasi K-Modes Clustering Untuk Pengelompokan Data Bermain Game Pada Mahasiswa Ditinjau Dari Durasi

- Belajarnya. *Media Online*), 4(5), 2495–2505.
<https://doi.org/10.30865/klik.v4i5.1619>
- Novidianto, R., & Fithriasari, K. (2021). Algoritma ClusterMix K-Prototypes Untuk Menangkap Karakteristik Pasien Berdasarkan Variabel Penciri Mortalitas Pasien Dengan Gagal Jantung. *Inferensi*, 4(1), 37.
<https://doi.org/10.12962/j27213862.v4i1.8479>
- Purba, Ariani Fitri; Mustafid, M., & Kartikasari, P. (2025). PENERAPAN ALGORITMA k-PROTOTYPE UNTUK PENGELOMPOKAN DESA DI KABUPATEN BEKASI BERDASARKAN INFRASTRUKTUR DIGITAL. 13, 479–489.
<https://doi.org/10.14710/j.gauss.13.2.479-489>
- Rahmawati, T., Wilandari, Y., & Kartikasari, P. (2024). Analisis Perbandingan Silhouette Coefficient Dan Metode Elbow Pada Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Ipm Dengan K-Medoids. *Jurnal Gaussian*, 13(1), 13–24.
<https://doi.org/10.14710/j.gauss.13.1.13-24>
- Ramdhan, W., Sitompul, O. S., Nababan, E. B., & Nasution, S. (2025). a Multi-View Clustering Approach and Exploratory Data Analysis for Dominant Demographic Knowledge Discovering on the Cause Behind the Pandemic. *ICIC Express Letters*, 19(2), 175–184.
<https://doi.org/10.24507/icicel.19.02.175>
- Reihanah, K. N., I Maruddani, D. A., & Widiharih, T. (2024). Clustering Karakteristik Industri Kecil Dan Menengah Di Kota Kendari Menggunakan Algoritma K-Prototypes. *Jurnal Gaussian*, 12(3), 340–351.
<https://doi.org/10.14710/j.gauss.12.3.340-351>
- Saniyah, N., Suarna, N., & Prihartono, W. (2024). Clustering Pelanggaran Lalu Lintas Pada Kendaraan Bermotor Menggunakan Algoritma K-Prototype (Studi Kasus: Pengadilan Negeri Cirebon). *Jurnal Mahasiswa Teknik Informatika*, 8(1), 772–778.
- Wita Setiawati, Sofi Mutiara Insani, & Syarifah Setiana Ardiati. (2024). Peran Marketing Terhadap Omset Penjualan UMKM. *MENAWAN: Jurnal Riset Dan Publikasi Ilmu Ekonomi*, 2(5), 52–61.
<https://doi.org/10.61132/menawan.v2i5.782>