
Pengelompokkan Kabupaten/Kota Provinsi Jawa Timur Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan Metode *Elbow* dan Algoritma *K-Prototype*

Mohamad Rivaldi Koni^{1*}, Ismail Djakaria², Nisky Imansyah Yahya³
^{1,2,3} Jurusan Matematika, Fakultas MIPA,
Universitas Negeri Gorontalo, Bone Bolango, 96554, Indonesia
* Corresponding author, email: mohamadrivaldikoni@gmail.com

Abstract

People's welfare is the goal of the State of Indonesia which is contained in the official state document, namely the opening of the 1945 Constitution paragraph IV, this can be interpreted to enjoy an affluent life, free from poverty and is a human right for every citizen in Indonesia. The grouping process is carried out to see the level of people's welfare for each Regency/City in East Java. In this study, before grouping, the number of clusters was selected using the Elbow method. After that do the grouping with the K-Prototype Algorithm. Furthermore, using Kruskal Wallis and Chi-Square, the test was carried out to determine the variables that influence grouping. The results of the study obtained the 3 best clusters using the Elbow method, grouping with the K-Prototype Algorithm where cluster 1 consisted of 4 Regencies/Cities, Cluster 2 consisted of 18 Regencies/Cities and cluster 3 consisted of 16 Regencies/Cities. Furthermore, the results of Kruskal Wallis and Chi-Square get 4 influential variables in the grouping, the 4 variables are the Number of Poor Population, Expenditures Per Capita, Open Unemployment Rate and Sources of Water for Drinking.

Keywords: Chi-Square, Elbow Method, K-Prototype Algorithm, Kruskal Wallis, People's Welfare

Abstrak

Kesejahteraan rakyat merupakan tujuan Negara Indonesia yang ada dalam dokumen resmi negara yaitu pembukaan UUD 1945 alinea IV, hal tersebut dapat diartikan menikmati kehidupan berkecukupan, bebas dari angka kemiskinan dan merupakan hak asasi bagi setiap warga negara di Indonesia. Proses pengelompokan dilakukan untuk melihat tingkat kesejahteraan rakyat untuk setiap Kabupaten/Kota di Jawa Timur. Dalam Penelitian ini sebelum melakukan pengelompokan, pemilihan jumlah klaster dilakukan terlebih dahulu dengan Metode Elbow. Setelah itu melakukan pengelompokan dengan Algoritma K-Prototype. Selanjutnya menggunakan *Kruskal Wallis* dan *Chi-Square*, uji tersebut dilakukan untuk mengetahui variable yang berpengaruh dalam pengelompokan. Hasil penelitian memperoleh 3 klaster terbaik menggunakan metode *Elbow*, pengelompokan dengan *Algoritma K-Prototype* dimana klaster 1 terdiri 4 Kabupaten/Kota, klaster 2 terdiri 18 Kabupaten/Kota dan klaster 3 terdiri 16 Kabupaten/Kota. Selanjutnya hasil *Kruskal Wallis* dan *Chi-Square* mendapatkan 4 variabel berpengaruh dalam pengelompokan, 4 variabel tersebut yaitu Jumlah Penduduk Miskin, Pengeluaran Perkapita, Tingkat Pengangguran Terbuka dan Sumber Air Untuk Minum.

Kata Kunci: Algoritma *K-Prototype*, *Chi-Square*, *Kruskal Wallis*, Kesejahteraan Rakyat, Metode *Elbow*

1. Pendahuluan

Analisis kluster adalah suatu analisis dalam ilmu statistik yang mempunyai tujuan untuk mengelompokkan suatu data atau objek yang dilihat dari kesamaan karakteristik diantara objek tersebut. Secara umum, analisis kluster terbagi menjadi dua analisis ,yaitu analisis kluster non kluster dan analisis kluster hirarki. Analisis kluster non hirarki merupakan pengelompokkan suatu data atau objek dimana banyaknya kelompok yang akan dibentuk ditentukan sebelum mengelompokkan suatu data atau objek [1]. Dalam analisis kluster non hirarki, ada beberapa metode yang selalu digunakan yaitu metode k-means dan metode k-modes. Metode k-means dikhususkan untuk tipe data numerik dan untuk metode k-modes dikhususkan untuk tipe data kategorikal [2]. Dalam dunia nyata, data lebih cenderung pada tipe data campuran sehingga ada modifikasi algoritma K-Modes dan K-Means, yaitu metode algoritma K-prototype.

K-prototype digunakan untuk objek data yang memiliki tipe data campuran yaitu numerik dan kategorik. Metode ini merupakan pengklasteran yang tepat untuk mengelompokkan dataset yang memiliki tipe data campuran. K-prototype menggunakan ukuran ketidaksamaan antar objek pada dataset. Algoritma ini menggunakan ukuran ketidaksamaan dengan menggabungkan jarak euclides dengan ukuran ketidaksamaan yang ada dalam k-modes [3]. Dalam K-prototype sama halnya dengan analisis kluster hirarki yaitu mengelompokkan data dimana banyaknya kluster dapat ditentukan terlebih dahulu. Untuk menentukan banyaknya kluster terbaik terdapat beberapa metode, diantaranya yaitu metode Elbow. Metode Elbow adalah salah satu metode dalam menentukan jumlah kluster yang tepat atau kluster yang optimal [3]. Cara menentukan kluster terbaik dari metode elbow yaitu dengan melihat penurunan inersia yang curam pada grafik sehingga membentuk sebuah siku, dimana Inersia merupakan hasil dari fungsi biaya yang sudah dihitung [4]. Dengan adanya penentuan kluster yang optimal, dapat memudahkan dalam menentukan jumlah kluster algoritma K-Prototypes. Metode ini dapat digunakan dalam berbagai permasalahan yang ada, diantaranya yaitu masalah tentang kesejahteraan rakyat.

Kesejahteraan rakyat merupakan tujuan dari Negara Indonesia yang ada dalam dokumen resmi negara yaitu pembukaan UUD 1945 alinea IV, hal tersebut dapat diartikan menikmati kehidupan yang berkecukupan, bebas dari angka kemiskinan dan itu merupakan hak asasi bagi setiap warga negara di Indonesia dan itu juga merupakan tugas bagi pemerintahan negara yang akan diwujudkan dengan adanya pembangunan ekonomi nasional [5]. Idealnya semakin besar tingkat pembangunan suatu negara, maka semakin meningkat pula kesejahteraan rakyat suatu negara. Kesejahteraan rakyat selalu akan menjadi salah satu topik yang memikat untuk di bahas demi meningkatkan tingkat kesejahteraan disetiap daerah di Indonesia terutama di Provinsi Jawa Timur.

Provinsi Jawa Timur merupakan salah satu provinsi di Indonesia yang harus melaksanakan tujuan negara Indonesia yaitu tentang kesejahteraan rakyat. Menurut Badan Pusat Statistika pada tahun 2021 jumlah penduduk di Provinsi Jawa Timur sebesar

40,67 juta jiwa dengan jumlah kepadatan penduduk sekitar 47,8 jiwa per kilometer persegi [6]. Jumlah penduduk Provinsi Jawa Timur mengalami peningkatan terus menerus. Ketika jumlah penduduk mengalami peningkatan terus menerus, berarti pemerintah juga harus melakukan atau menambah jumlah fasilitas kehidupan untuk memenuhi kebutuhan layak hidup bagi masyarakat. Ditambah lagi dengan adanya tingkat pengangguran yang semakin bertambah. Tingkat pengangguran di Provinsi Jawa Timur terakhir menurut Badan Pusat Statistik yaitu sebesar 9,54 % pada bulan Agustus 2021. Jika tidak ada perubahan strategi dalam hal untuk mengatasi tingkat pengangguran tersebut akibatnya tingkat pengangguran akan semakin meningkat setiap waktu. Maka dari itu solusi yang tepat untuk masalah tersebut yaitu dengan cara mengidentifikasi karakter-karakter berdasarkan indikator kesejahteraan rakyat, disini pemerintah dapat memutuskan kebijakan dan strategi yang tepat dalam pembangunan untuk memenuhi kebutuhan yang layak bagi rakyat.

Pada penelitian sebelumnya yang dilakukan oleh Winarta dan Kurniawan pada tahun 2021 tentang Optimasi cluster K-Means menggunakan Metode Elbow Pada Data Pengguna narkoba dengan pemrograman Python [7]. Pada Penelitian ini, metode K-means dikombinasikan dengan metode Elbow untuk mengelompokkan data pengguna Narkoba, perbedaan antara penelitian ini dan penelitian penulis yaitu pada metode dan studi kasusnya dimana penulis menggunakan metode Elbow yang dikombinasikan dengan Algoritma K-prototype untuk mengelompokkan Kabupaten/Kota berdasarkan indikator kesejahteraan rakyat. Selanjutnya pada tahun 2018 penelitian yang dilakukan oleh Alwi dan Hasrul dengan judul Analisis kluster untuk pengelompokkan Kabupaten/Kota di provinsi Sulawesi Selatan berdasarkan indikator kesejahteraan rakyat [1]. Pada Penelitian ini, peneliti menggunakan topik yang sama dengan penelitian tersebut yaitu indikator Kesejahteraan Rakyat tetapi menggunakan metode yang beda dari penelitian tersebut. Selanjutnya yaitu penelitian yang dilakukan oleh Andrew Donda Munthe dkk pada tahun 2018 tentang Penggrerombolan Desa/Kelurahan Berdasarkan Indikator Kemiskinan Dengan Algoritma TSC dan K-Prototypes [8]. Pada penelitian tersebut, peneliti menggunakan metode yang sama yaitu Algoritma K-Prototypes tetapi dikombinasikan dengan metode Elbow dalam pemilihan kluster terbaik.

Berdasarkan penjelasan yang telah diuraikan, peneliti tertarik mengambil atau melakukan penelitian dengan judul “Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur berdasarkan indikator kesejahteraan rakyat menggunakan kombinasi metode Elbow dan Algoritma K-Prototype”.

2. Material dan Metode

2.1 Sumber Data

Data yang dikumpulkan pada penelitian ini adalah data sekunder hasil dari pendataan Badan Pusat Statistika (BPS) pada tahun 2021. Sampel dalam penelitian adalah seluruh Kabupaten/Kota di Provisi Jawa Timur dengan total sebanyak 38 Kabupaten/Kota.

Variabel yang digunakan sebanyak 7 variabel (3 variabel numerik dan 4 variabel kategorik). Variabel-variabel tersebut dipilih karena termasuk dalam perubahan taraf kesejahteraan rakyat menurut Badan Pusat Statistik (BPS). Variabel-variabel yang akan digunakan pada penelitian ini sebagai berikut :

- a. Jumlah penduduk miskin (X_1)
- b. Pengeluaran perkapita (X_2)
- c. Tingkat pengangguran terbuka (X_3)
- d. Topografi wilayah (X_4)
- e. Kepemilikan tempat tinggal sebagian besar keluarga (X_5)
- f. Tempat buang air besar sebagian besar keluarga (X_6)
- g. Sumber air untuk minum sebagian besar keluarga (X_7)

2.2 Metode *Elbow*

Metode *Elbow* atau siku adalah metode penentuan jumlah kelompok yang tepat melalui persentasi hasil perbandingan antara jumlah Klaster satu dengan jumlah klaster yang lain. Metode *Elbow* memberikan gambaran menggunakan skor klaster dan menambah skor klaster sebagai model data dalam menentukan klaster terbaik. Optimasi nilai klaster (k) akan didapatkan dengan biaya yang digunakan [9]. Berikut merupakan definisi rumus fungsi dari biaya yang digunakan dalam algoritma K-Prototype :

$$SSE = \sum_{l=1}^k \sum_{i=1}^n ||(x_i - u_k)||^2 \quad (1)$$

dengan,

x_i = Nilai atribut dari data ke i

u_k = Nilai atribut titik pusat klaster i

2.3 Algoritma *K-Prototype*

Algoritma *K-prototype* merupakan pengintegrasian dari metode *K-means* dan *K-modes*. Secara efektif akan lebih berguna dikarenakan banyak ditemui dalam dunia nyata adalah objek tipe data gabungan dari numerik dan katedorikal atau dengan kata lain data campuran. *K-Prototype* juga menggunakan ketidaksamaan antara objek dan data set. Ukuran ketidaksamaan yang akan digunakan juga berbeda dengan K-Modes. Untuk mengukur ketidaksamaan (*dissimilarity measure*) yaitu menggabungkan euclidien distance dengan dissimilarity measure didalam *K-Modes* [10]. Berikut merupakan ukuran jarak yang dilakukan untuk algoritma *K-Prototype*.

$$d_2(X, Y) = \sum_{i=1}^p (x_j - y_j) + \gamma \sum_{p+1}^m \delta(x_j, y_j) \quad (2)$$

dengan,

D_2 : Merupakan symbol untuk *dissimilarity measure*

γ : Hasil yang didapat setelah proses konvergen

2.4 Uji Beda

Uji beda ini dilakukan untuk mengetahui variabel mana yang berpengaruh atas pengelompokan data atau variabel mana yang lebih dominan dalam hal pengelompokan data. Terdapat beberapa metode dalam melakukan uji beda diantaranya yaitu uji *Kruskal Wallis* dan uji *Chi-Square*. Uji *Kruskal Wallis* merupakan uji yang akan digunakan untuk melihat perbedaan signifikan antara dua atau lebih kelompok bebas jika skala data untuk variabel terikat yaitu ordinal atau interval/ratio. Untuk mengetahui pengaruh dua kelompok atau lebih kelompok data itu dilihat dari nilai p-value nya. Jika Nilai P-Value $> 0,05$ maka terima H_0 begitupun sebaliknya [11]. Sedangkan untuk uji *Chi-Square* atau biasa disebut dengan *Chi-Kuadrat* merupakan uji non parametrik yang prinsip kerjanya adalah dengan membandingkan dua kelompok data yang skala datanya adalah nominal. Untuk mengetahui pengaruh 2 kelompok data sama seperti *Kruskal Wallis* yaitu jika *P-Value* $> 0,05$ maka terima H_0 begitupun sebaliknya [12].

3. Hasil dan Diskusi

3.1 Deskripsi variabel.

Analisis deskriptif dilakukan untuk masing-masing variabel yang akan digunakan dalam penelitian, untuk variabel numerik meliputi : jumlah penduduk miskin, pengeluaran perkapita dan tingkat pengangguran terbuka. Nilai maksimal dari variabel diatas yaitu 276.560, 17 862.000 dan 10.870 dengan nilai rata-rata 120.335, 11.568.526, dan 5.520. Variabel kategorik meliputi : topografi wilayah, kepemilikan tempat tinggal sebagian besar keluarga, tempat buang air besar sebagian besar keluarga dan sumber air untuk minum sebagian besar keluarga. Berikut merupakan deskripsi variable numerik maupun kategorik :

Tabel 1. Variabel Numerik

Variabel	Keterangan	Maksimal	Rata-rata
X ₁	Jumlah penduduk miskin	276.569	120.335
X ₂	Pengeluaran perkapita	17.862.000	11.568526
X ₃	Tingkat pengangguran terbuka	10.870	5.520

Sumber : Hasil Pengolahan

Tabel 2. variabel Kategorik

Variabel	Keterangan	Kategorik
X ₄	Topografi wilayah	2
X ₅	Kepemilikan tempat tinggal sebagian besar keluarga	2

X ₆	Tempat buang air besar sebagian besar keluarga	1
X ₇	Sumber air untuk minum sebagian besar keluarga	2

Sumber : Hasil Pengolahan

3.2 Pre Processing

Data yang berhasil terkumpul terdiri dari 38 baris dan 7 kolom. Data yang terkumpul tersebut harus diubah menjadi format yang lebih mudah dan lebih efektif agar dalam proses pengelompokan berjalan dengan lancar. Caranya yaitu dengan cara mengkonversi datanya ke dalam bentuk matriks. Konversi data dalam bentuk matriks ini dilakukan agar proses pengelompokan data berjalan dengan lancar sesuai dengan tujuan penelitian yang akan dilakukan.

3.3 Metode *Elbow*

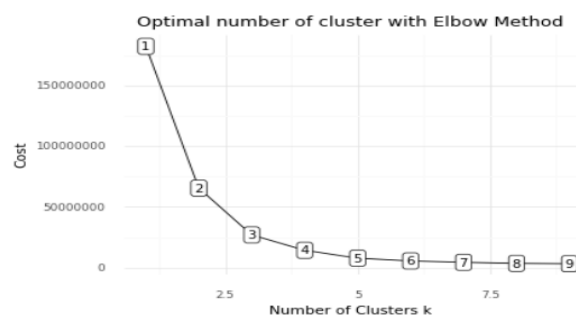
Berdasarkan rumus pada persamaan 1, hasil pengujiannya dapat dilihat pada tabel 3 berikut ini.

Tabel. 3. Pengujian Metode *Elbow*

Klaster	Cost
1	182101574.995
2	65509680
3	27017210.372
4	14376905.087
5	7869407.613

Sumber : Hasil Pengolahan

Berdasarkan hasil pengujian dengan menggunakan metode *Elbow* menunjukkan bahwa terdapat perbedaan yang signifikan antara klaster 1 dan 2 dan juga klaster 2 dan 4, dan untuk klaster selanjutnya tidak terdapat perbedaan yang sangat signifikan. Cara menentukan klaster yang optimal dapat dilihat pada titik yang membentuk siku pada gambar 1 sebagai berikut.

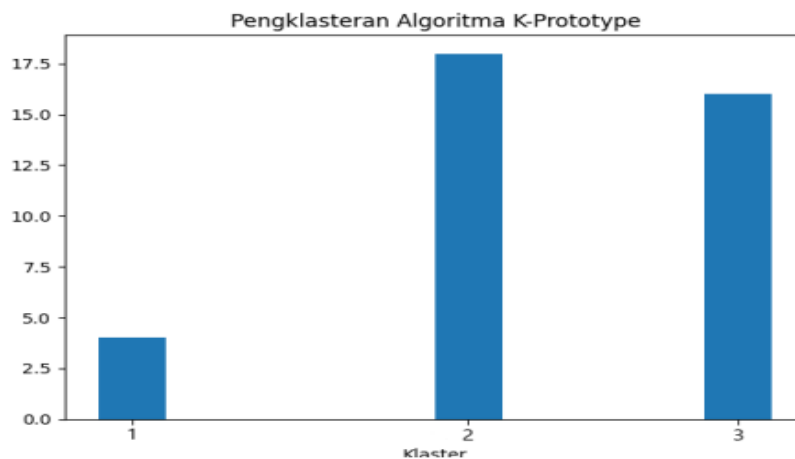


Gambar 1. Grafik Metode *Elbow*

Berdasarkan gambar 1, dapat dilihat bahwa titik yang membentuk siku itu diantara klaster 2 atau klaster 3. Dapat kita ketahui bersama bahwa maksud dari titik yang membentuk siku yaitu penurunan antara 2 titik yang sangat signifikan selanjutnya diikuti oleh titik klaster yang relatif konstan. Berdasarkan maksud siku yang sudah dijelaskan berarti klaster optimal yaitu 3 klaster karena terdapat penurunan yang sangat signifikan antara klaster 2 dan 3 dan diikuti oleh klaster yang relatif konstan.

3.4 Algoritma *K-Prototype*

Langkah awal dalam pengklasteran untuk *K-Prototype* yaitu menentukan jumlah kelompok. Jumlah kelompok dipilih berdasarkan pengujian dengan metode penentuan Jumlah klster yang optimal yaitu metode *Elbow*. Dimana jumlah yang telah ditentukan terdiri dari 3 klaster optimal. Hasil dari proses pengklasteran dapat dilihat dalam gambar 2 berikut :



Gambar 2. Hasil K-Prototype

3.5 Uji Beda Hasil Klaster

Tahap ini merupakan tahap untuk menentukan variabel apa yang lebih dominan dalam pengelompokan data. Pada tahap ini penulis menggunakan uji *Kruskal Wallis* untuk data numerik, sedangkan untuk tipe data kategorik menggunakan uji *Chi-Square* dalam menentukan variabel yang lebih dominan dalam pengelompokan Kabupaten/Kota. Untuk uji *Kruskal Wallis* dan uji *Chi-Square* jika nilai *p-value* < 0,05 maka Terima H0 begitupun sebaliknya, jika *p-value* > 0,05 maka tolak H0. Berikut merupakan hasil perhitungan dari uji beda hasil klaster :

Tabel 4. *Kruskal Wallis*

Variabel	P-value	Keputusan
X ₁	1.38513e-06	Terima H ₀
X ₂	0.00032	Terima H ₀

X ₃	0.00285	Terima H ₀
----------------	---------	-----------------------

Sumber : Hasil Pengolahan

Tabel 5. *Chi-Square*

Variabel	P-value	Keputusan
X ₄	0,788128	Tolak H ₀
X ₅	1	Tolak H ₀
X ₆	1	Tolak H ₀
X ₇	0.00673795	Terima H ₀

Sumber : Hasil Pengolahan

Berdasarkan tabel 4 dan tabel 5, Uji beda antar klaster pada variabel X₁, X₂, X₃ dan X₇ memperoleh nilai *p-value* < 0,05. sehingganya terima H₀, artinya terdapat perbedaan yang signifikan atau dapat dikatakan memiliki perbedaan yang signifikan antar klaster. Untuk variabel lainnya itu memperoleh nilai *p-value* > 0,05 sehingganya tolak H₀, artinya tidak terdapat perbedaan yang signifikan atau dapat dikatakan tidak memiliki perbedaan yang signifikan antar klaster.

Berdasarkan hasil uji *Kruskal Wallis* dan uji *Chi Square* Variabel X₁ (Jumlah penduduk miskin), X₂ (Pengeluaran Perkapita), X₃ (Tingkat Pengangguran Terbuka) dan X₇ (Sumber air Untuk Minum Sebagian Besar Keluarga) yang memiliki perbedaan yang signifikan antar klaster, karena variabel X₁ (Jumlah Penduduk Miskin), X₂ (Pengeluaran Perkapita), X₃ (Tingkat Pengangguran Terbuka) dan X₇ (Sumber Air Minum) Hasil *p-value* nya < 0,05. sehingganya Jumlah Penduduk miskin, Pengeluaran Perkapita, Tingkat Pengangguran Terbuka, dan Sumber air minum merupakan variabel Pembeda antar Klaster yang terbentuk atau merupakan Variabel yang lebih dominan dalam pengelompokan data.

4. Kesimpulan

Berdasarkan Hasil Penelitian yang telah dilakukan dapat disimpulkan bahwa:

1. Pada Penentuan jumlah klaster terbaik menggunakan metode *Elbow*, klaster yang didapat yaitu 3 klaster, klaster tersebut didapat karena nilai klaster 2 (65509680.872) berbeda jauh dengan klaster 3 (270272210.372) serta diikuti oleh klaster lainnya yang nilainya relatif konstan. maka dari itu klaster optimal yang telah ditetapkan yaitu 3 klaster.
2. Pengelompokan Kabupaten/Kota berdasarkan indikator kesejahteraan rakyat dengan menggunakan Algoritma *K-prototype* membentuk tiga klaster. Penentuan klaster tersebut dilakukan dengan Pengujian metode *Elbow*. Variabel yang lebih dominan dalam hal pengelompokan Kabupaten/Kota yaitu Jumlah Penduduk Miskin (X₁), Pengeluaran Perkapita (X₂), Tingkat Pengangguran Terbuka (X₃) dan juga Sumber Air Untuk Minum Sebagian Besar Keluarga (X₇). Adapun hasil

pengelompokan Kabupaten/Kota berdasarkan indikator Kesejahteraan rakyat yaitu 4 Kabupaten/Kota termasuk pada klaster satu yang terdiri dari Sidoarjo, Kota Malang, Kota Madiun, Kota Surabaya. 18 Kabupaten/Kota lainnya termasuk pada klaster dua yaitu Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Malang, Lumajang, Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Bojonegoro, Tuban, Bangkalan, Sampang, Pamekasan, Sumenep, dan 16 Kabupaten/Kota termasuk pada klaster 3 yang terdiri dari Kediri, Banyuwangi, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi.

Daftar Pustaka

- [1] W. Alwi dan M. Hasrul, “Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat,” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 6, no. 1, hal. 35, 2018, doi: 10.24252/msa.v6i1.4782.
- [2] A. Badruttamam, S. Sudarno, dan D. A. I. Maruddani, “Penerapan Analisis Klaster K-Modes dengan Validasi Davies Bouldin Index dalam Menentukan Karakteristik Kanal Youtube di Indonesia (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade),” *J. Gaussian*, vol. 9, no. 3, hal. 263–272, 2020, doi: 10.14710/j.gauss.v9i3.28907.
- [3] A. S. Sulthoni, R. Andreswari, dan F. Hamami, “Segmentasi Pelanggan Pt. Telekomunikasi Seluler Indonesia Menggunakan Clustering Algoritma K-prototypes dan Metode Elbow Sebagai Perumusan Strategi Marketing,” *eProceedings Eng.*, vol. 8, no. 3, 2021.
- [4] A. F. Febrianti, A. H. Cabral, dan G. Anuraga, “K-Means Clustering Dengan Metode Elbow untuk Pengelompokan Kabupaten dan Kota di Jawa Timur,” *Semin. Nas. Has. Ris. dan Pengabd. -SNHRP*, hal. 863–870, 2018, [Daring]. Tersedia pada: <http://karyailmiah.unipasby.ac.id/wp-content/uploads/2019/04/K-Means-Artikel.pdf>
- [5] D. Nawarti Bustaman, “Analisis Faktor-Faktor yang Mempengaruhi Indikator Kesejahteraan Masyarakat di Kota Pekanbaru,” *Ekon. Kiat*, vol. 32, no. 1, hal. 85–92, 2021, [Daring]. Tersedia pada: <https://journal.uir.ac.id/index.php/kiat/article/view/7677/3567>
- [6] “Badan Pusat Statistik,” 2021.
- [7] A. Winarta dan W. J. Kurniawan, “Optimasi cluster k-means menggunakan metode elbow pada data pengguna narkoba dengan pemrograman python,” *J. Tek. Inform. Kaputama*, vol. 5, no. 1, hal. 113–119, 2021.
- [8] A. D. Munthe, I. M. Sumertajaya, dan U. D. Syafitri, “Penggerombolan Desa/Kelurahan Berdasarkan Indikator Kemiskinan Dengan Menerapkan Algoritma Tsc dan K-Prototypes,” 2018. doi: 10.29244/ijsa.v2i2.169.
- [9] D. A. I. C. Dewi dan D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi

- Kerajinan Bali,” *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, hal. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [10] R. Nooraeni, S. Tinggi, dan I. Statistik, “Metode Cluster Menggunakan Kombinasi Algoritma Cluster K-Prototype Dan Algoritma Genetika Untuk Data Bertipe Campuran Cluster Method Using a Combination of Cluster K-Prototype Algorithm and Genetic Algorithm for Mixed Data,” *J. Apl. Stat. Komputasi Stat.*, vol. 7, no. 2, hal. 17–17, 2015, [Daring]. Tersedia pada: <https://jurnal.stis.ac.id/index.php/jurnalasks/article/view/23>
- [11] A. Mubarok dan Sa. Susanto, “Mann Whitney Test in Comparing the Students’ Consultation Results of Entrepreneurial Practice Between Male and Female Lecturers in Economic Faculty of Pamulang University,” *Procur. J. Ilm. Manaj. Procur.*, vol. 9, no. 1, hal. 9–15, 2021.
- [12] A. Wibowo, “Uji Chi-Square pada Statistika dan SPSS,” *J. Ilm. Sinus*, vol. 4, no. 2, hal. 38, 2017.