
Implementasi *Clustering K-Medoids* dalam Pengelompokan Kabupaten di Provinsi Aceh Berdasarkan Faktor yang Mempengaruhi Kemiskinan

Freditasari Purwa Hidayat*, Royhan Pina Putra, M Dendi Alfitrah, Edy Widodo
Universitas Islam Indonesia

**Corresponding author:* 18611027@students.uii.ac.id

Abstract. *The economy is one of the parameters to see how the development of a country. Ending poverty anywhere and in any form is goal 01 of the Sustainable Development Goals (SDGs) program. Until now, poverty has become one of the main problems in Indonesia, so poverty must be a concern of the government. Based on data from the Central Statistics Agency (BPS) shows that as of September 2020 the percentage of poor people in Aceh Province is still the highest on the island of Sumatra, which is 15.43%. The purpose of this study is to classify districts based on factors that affect poverty in Aceh Province. The method used in this study is the K-Medoids Cluster Analysis algorithm. The optimal number of clusters is 2 clusters with cluster 1 consisting of 11 districts and cluster 2 consisting of 12 districts. Cluster 1 has a higher percentage of poor population and poverty depth index than cluster 2, while cluster 2 has higher Gini Ratio, AHH, and RLS values than cluster 1.*

Keywords : *Clusters, Economy, Poverty, SDGs*

1. PENDAHULUAN

Kemiskinan merupakan suatu keadaan dimana ketidakcukupan seseorang dalam mencukupi kebutuhan primer seperti sandang, pangan, papan, pendidikan, maupun kesehatan. Kemiskinan merupakan permasalahan yang kompleks dan selalu menjadi momok bagi Negara berkembang salah satunya adalah Negara Indonesia, dimana Indonesia yang berada pada peringkat ke-5 negara termiskin di Asia Tenggara. Kemiskinan adalah permasalahan kompleks dan bersifat multidimensional yang membuat pusat perhatian bagi pemerintah di seluruh dunia. Permasalahan kemiskinan dapat mempengaruhi dalam banyak aspek seperti sosial, ekonomi, budaya, serta lainnya. Salah satu parameter keberhasilan pembangunan adalah pertumbuhan ekonomi yang tinggi [1]. Oleh sebab itu, diperlukan cara yang dapat mengurangi tingkat kemiskinan di Indonesia salah satu caranya yaitu meningkatkan perekonomian.

Tujuan pertama dari *Sustainable Development Goals* (SDGs) adalah “Mengakhiri Kemiskinan dalam Segala Bentuk Di Mana Pun” [2]. Adapun penelitian yang dikerjakan oleh Tri Santoso menyatakan bahwa pengentasan kemiskinan mempunyai keterkaitan dengan tujuan lain yang diumumkan oleh SDGs [3].

Dari data BPS menunjukkan bahwa sampai dengan September 2020 persentase penduduk miskin di Provinsi Aceh masih menjadi yang tertinggi di Pulau Sumatera yaitu sebesar 15,43%. Di bawah Provinsi Aceh terdapat Provinsi Bengkulu dengan persentase penduduk miskinnya

sebesar 15,3%, Provinsi Sumatera Selatan sebesar 12,98%, Provinsi Lampung sebesar 12,76%, Provinsi Sumatera Utara sebesar 9,14%, Provinsi Jambi sebesar 7,97%, Provinsi Riau sebesar 7,04%, Provinsi Kepulauan Riau sebesar 6,13% dan Provinsi Bangka Belitung sebesar 4,89%.

Dalam melakukan penelitian, sangat diperlukan kajian mengenai penelitian terdahulu. Persoalan ini sangat berguna sebagai acuan agar penelitian yang akan dibuat dapat menjadi lebih baik daripada penelitian sebelumnya. Kajian penelitian yang dilakukan Disty Wahyuli, 2019 berjudul Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma *K-Medoids* [4]. Dari hasil *K-Medoids* yang dilakukan dalam kajian penelitian ini didapatkan kesimpulan bahwa terdiri dari 11 provinsi yang memiliki tingkat kemiskinan rendah, sedangkan 23 provinsi dengan tingkat kemiskinan rendah.

Berdasarkan poin-poin yang telah dijelaskan peneliti sebelumnya, peneliti hendak melakukan pengelompokan kabupaten/kota di Provinsi Aceh berdasarkan faktor-faktor yang mempengaruhi kemiskinan di Provinsi Aceh. Pengelompokan ini dilakukan menggunakan analisis *clustering K-Medoids*. Hasil pengelompokan dapat dijadikan acuan bagi pemerintah dalam mengatasi permasalahan kemiskinan di Provinsi Aceh. Selain itu, dengan dibentuknya kelompok daerah tersebut dapat mempermudah pemerintah dalam menjalankan rencana/program kerja yang akan ditargetkan di setiap kabupaten/kota Provinsi Aceh.

Namun, penelitian ini memiliki beberapa perbedaan dengan kajian pustaka yang digunakan. Perbedaan tersebut yaitu pada penelitian ini dilakukan uji asumsi terlebih dahulu sebelum melakukan *clustering* menggunakan metode *K-Medoids*, peneliti menggunakan lebih banyak variabel dalam melakukan *clustering* menggunakan metode *K-Medoids*, dan adanya visualisasi peta yang mempermudah tersampainya informasi menggunakan *software GIS*.

2. LANDASAN TEORI

2.1 Kemiskinan

Mengakhiri segala bentuk kemiskinan merupakan tujuan SDGs yang pertama. Di Indonesia kemiskinan merupakan persoalan yang sampai sekarang belum terselesaikan karena adanya kesenjangan sosial, Kondisi orang yang mengalami kemiskinan dilihat dari kemampuan untuk mencukupi kehidupannya [5]. Keadaan ini ditandai dengan pendapatan yang tidak mencukupi untuk melengkapi kebutuhan dasar seperti sandang, pangan, dan papan. Indonesia sebagai negara yang merupakan anggota dari PBB berjanji untuk menyelesaikan masalah yang sesuai dengan deklarasi SDGs. Artinya Indonesia juga perlu mencapai tujuan yang telah ditetapkan dalam dsiaran PBB [6].

2.2 Sustainable Development Goals (SDGs)

Banyaknya masalah yang sering terjadi sampai sekarang menjadi sebuah tanggung jawab setiap negara di dunia. Mulai dari kemiskinan, hidup sehat, energi yang menipis, kesetaraan gender, pencemaran lingkungan, dan lain sebagainya. Dunia meresponnya dengan menyetujui deklarasi yang dikenal dengan *Millennium Development Goals* (MDGs). Dengan berakhirnya MDGs pada tahun 2015, maka dilakukan pengembangan konsep dalam agenda pembangunan pasca-2015 yang disebut dengan *Sustainable Development Goals* (SDGs). Terdapat 17 sudut pandang tujuan pembangunan berkelanjutan yang berisi 169 tujuan yang telah dirumuskan dari hasil sidang PBB [2].

2.3 Analisis Deskriptif

Terdapat 2 macam metode statistik, yaitu statistik deskriptif dan inferensial. Statistika deskriptif ialah metode yang dipakai untuk menampilkan gambaran data hasil penelitian akan tetapi tidak sampai hingga menarik kesimpulan yang lebih mendalam terhadap sifat populasi (generalisasi/inferensi). Ruang lingkup dari statistika deskriptif mencakup: konsep dasar statistika, distribusi frekuensi, pengukuran nilai pusat (*central tendency*), pengukuran penyebaran (*dispersion*), kemiringan (*skewness*) dan keruncingan (*kurtosis*), penyajian data dalam bentuk diagram grafik (diagram garis, diagram batang, batang histogram, *polygon*, *ogive*), angka indeks, dan *time series* atau deret waktu [7].

2.4 Uji Multikolinearitas

Terdapat beberapa cara untuk melihat multikolinearitas, yaitu tidak terjadi gejala multikolinearitas apabila korelasi antar variabel independen bernilai < 0.95 , pengujian ini dapat dilihat pada nilai koefisien korelasi pearson [8]. Selain itu, dapat dilihat pada nilai *Variance Inflation Factor (VIF)*, tidak terbentuk multikolinearitas jika *VIF*-nya bernilai di bawah 10 [9].

2.5 Analisis Cluster

Analisis *cluster* adalah analisis untuk menggabungkan sejumlah objek ke dalam satu kelompok yang relatif sama. *Cluster* yang terbentuk cenderung memiliki kesamaan objek, namun berbeda jauh antar *cluster* yang lain [10]. Tidak seperti halnya pada klasifikasi, variabel target tidak terdapat pada analisis *cluster*. Menurut [11], variabel target tidak diklasifikasi, diramal, atau diprediksi pada *clustering*.

Pada analisis *cluster* terdapat 2 metode yang bisa dipakai yaitu *hierarchical clustering* dan *non-hierarchical clustering*. Terdapat perbedaan dari kedua metode dapat dilihat pada Tabel 1.

Tabel 1. Perbedaan Metode *Clustering* Hierarki dan Non Hierarki

Metode Hierarki	Metode Non Hierarki
Dapat digunakan untuk data yang tidak terlalu banyak [12].	Dapat diterapkan pada data yang lebih besar dibandingkan dengan metode hierarki.
Jumlah kelompok yang hendak dibentuk belum diketahui [12].	Jumlah kelompoknya ditentukan terlebih dahulu [13].
Beberapa metode pengklasteran hierarki yaitu <i>single linkage</i> , <i>complete linkage</i> , <i>average linkage</i> , dan <i>ward's method</i> .	Metode pengelompokan sekatan pada penelitian ini yaitu <i>K-Medoids Clustering</i> .

2.6 K-Medoids Clustering

Metode ini merupakan salah satu metode analisis statistik untuk mengelompokan data. Perbedaan mendasar metode ini dibandingkan dengan *K-Means Clustering* yaitu pada perhitungan *mean* dan jarak dari data berdasarkan tiap-tiap *mean*. Sedangkan pada *K-Medoids Clustering*, data dipakai sebagai nilai tengah (*medoid*), lalu perhitungan jarak yang dihitung dari selisih antar tiap data [14]. *K-Medoids Clustering* dapat digunakan untuk menanggulangi kekurangan *K-Means Clustering* yang sensitif terhadap data pencilan (*outlier*) [15]. Dalam menggabungkan sekumpulan n objek menjadi sebanyak k kluster digunakan metode pengklasteran partisi pada *K-Medoids Clustering*. Algoritma yang digunakan adalah *medoids*, yaitu objek pada sekelompok objek yang mewakili sebuah kluster. Kedekatan antara objek

medoids dengan objek *non medoids* dihitung untuk membangun sebuah *cluster* [16]. Langkah-langkah untuk menggunakan algoritma *k-medoids* [16], yaitu:

1. Menentukan jumlah *cluster* yang dibentuk dengan pertimbangan teoritis dan konseptual.
2. Persamaan ukuran jarak antara *Euclidean distance*, untuk menghitung setiap objek terhadap *cluster* terdekat.

$$d_{ab} = \sqrt{\sum_{n=1}^p (x_{an} - x_{bn})^2} \tag{1}$$

Keterangan:

d_{ab} = jarak antara objek a dengan b

x_{an} = nilai pada variabel ke-n

x_{bn} = nilai pada variabel ke-n

p = banyaknya variabel yang diamati

3. Kemudian menghitung jarak *Euclidean distance* pada setiap objek untuk calon *non medoids* dengan menentukan pusat *cluster* baru secara random.
4. Selanjutnya, melakukan perhitungan jarak pada objek-objek yang terletak pada *cluster-cluster* calon *non medoids*.
5. Setelah itu, melakukan perhitungan antara total simpangan (S) dengan total jarak baru – total jarak lama. Ketika membentuk sekelompok n objek baru sebagai *medoids* pada waktu jika $S < 0$ lalu tukar objek dengan data *cluster non medoids*.

Ulangilah cara 3 – 5 hingga tidak terdapat perubahan pada *medoids*, maka diperoleh *cluster* dan anggota setiap *cluster*.

2.7 Interpretasi Profil Cluster

Interpretasi hasil ialah kegiatan yang paling akhir dalam analisis *cluster* yang dilakukan dengan cara memeriksa variasi *cluster* setiap kelompok guna memberi label agar dapat menggambarkan karakter dari sebuah *cluster* secara tepat. Nilai *centroid cluster* merupakan parameter yang sering dipakai ketika melakukan interpretasi [16]. Profilisasi dilakukan guna menjelaskan perbedaan antar *cluster* dengan cara penggambaran karakteristik dari masing-masing *cluster*.

3. METODE PENELITIAN

Pada penelitian ini menggunakan data sekunder. Data tersebut adalah data ekonomi yang didapatkan dari *website* Badan Pusat Statistika Provinsi Aceh (URL: <https://aceh.bps.go.id/>). Dalam penelitian ini menggunakan 5 variabel yaitu variabel Persentase Penduduk Miskin, Angka Harapan Hidup (AHH), Rata-Rata Lama Sekolah (RLS), Gini Ratio, dan Indeks Kedalaman Kemiskinan (IKK) di Provinsi Aceh tahun 2020.



Gambar 1. Flowchart Penelitian

Berikut merupakan langkah-langkah penelitian yang akan dilakukan oleh peneliti:

1. Memasukan data kemiskinan Provinsi Aceh dengan variabel persentase penduduk miskin, RLS, AHH, gini ratio, dan IKK.
2. Melakukan analisis deskriptif menggunakan *software R* untuk mengetahui nilai *min*, *max*, dan *mean*. Analisis deskriptif bisa digunakan untuk melihat apakah terdapat *outlier* yang berada pada data.
3. Selanjutnya, melakukan uji multikolinearitas untuk mengetahui terjadinya multikolinieritas atau tidak.
4. Melakukan pengecekan *outlier*.
5. Menentukan jumlah *cluster* yang optimum menggunakan metode yaitu *sillhouette*.
6. Melakukan analisis dan interpretasi dari hasil *cluster* berdasarkan metode *clustering K-Medoids* dengan perhitungan jarak *Euclidean*.
7. Menyimpulkan hasil pengelompokan *cluster*, melihat karakteristik dari masing - masing *cluster*.
8. Melakukan visualisasi hasil *cluster* menggunakan peta.

4. HASIL DAN PEMBAHASAN

4.1 Analisis Deskriptif

Berikut merupakan gambaran umum faktor-faktor yang mempengaruhi kemiskinan tahun 2020 di Provinsi Aceh.

Tabel 2. Analisis Deskriptif

Hasil	Variabel				
	Persentase Penduduk Miskin (%)	Gini Ratio (%)	IKK (%)	AHH (%)	RLS (%)
<i>Min</i>	6,90	0,2270	1,000	64,02	7,84
<i>Mean</i>	15,35	0,2915	2,371	68,27	9,43
<i>Max</i>	20,20	0,3560	3,550	71,60	12,65

Berdasarkan pada Tabel 2 didapatkan bahwa variabel AHH memiliki nilai minimum, rata-rata, dan nilai maksimum tertinggi yaitu nilai minimum sebesar 64,02, rata-rata sebesar 68,27, dan nilai maksimum sebesar 71,6. Suatu kabupaten/kota dikatakan memiliki tingkat kemiskinan rendah apabila memiliki nilai variabel Persentase Penduduk Miskin, gini ratio, dan IKK yang rendah tetapi memiliki nilai variabel AHH dan RLS yang tinggi.

3.1 Uji Multikolinearitas

Berikut merupakan nilai VIF dari masing-masing variabel yang digunakan dalam penelitian ini.

Tabel 3. Nilai VIF Dari Masing-Masing Variabel

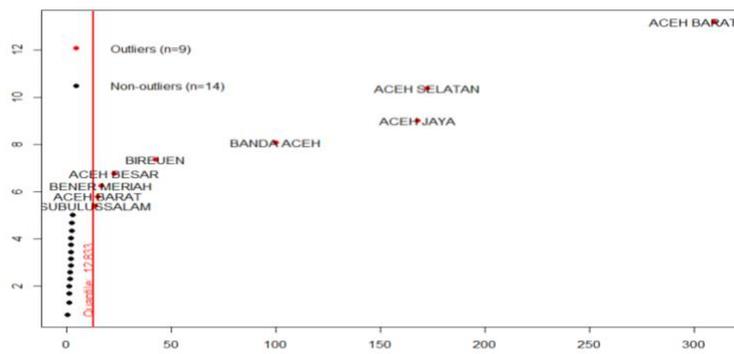
Hasil <i>VIF</i>	Variabel
3,402839	Persentase Penduduk Miskin
1,397099	Gini Ratio
2,510584	IKK
2,077331	AHH
2,553056	RLS

Dari Tabel 3 dapat dilakukan pengujian hipotesis multikolinearitas dengan melihat nilai VIF untuk masing-masing variabel sebagai berikut:

1. Hipotesis
 - H_0 : Tidak terjadi multikolinearitas
 - H_1 : Terjadi multikolinearitas
2. Tingkat Signifikansi
 - $\alpha = 5\% = 0,05$.
3. Daerah Kritis
 - Jika $VIF > 10$, maka tolak H_0
4. Statistik Uji dan Keputusan
 - Gagal Tolak H_0 , karena tidak ada variabel terdapat nilai $VIF > 10$
5. Kesimpulan
 - Dengan tingkat signifikansi sebesar 5% atau tingkat kepercayaan 95% didapatkan keputusan tolak H_0 untuk semua variabel, maka diperoleh kesimpulan bahwa tidak terdapat korelasi antar variabelnya (tidak terjadi multikolinearitas) sehingga asumsi terpenuhi.

3.2 Pengecekan *Outlier*

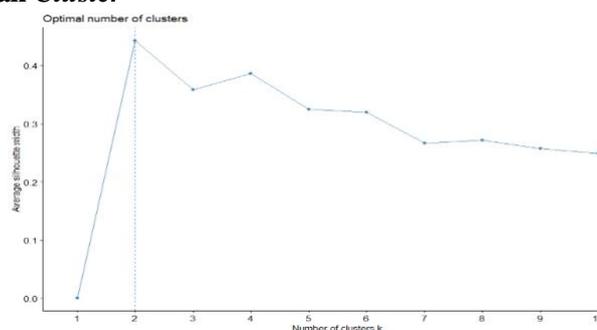
Analisis *K-Medoids clustering* digunakan untuk mengelompokkan data yang mengandung nilai *outlier*, oleh karena itu dilakukan pengecekan *outlier* dengan metode quan sebagai berikut.



Gambar 2. Pengecekan *Outlier*

Dari *output* yang ada pada Gambar 2 dapat dilihat bahwa dengan metode quan ditemukan adanya 9 data *outlier* dan 14 data *non-outlier*. Data *outlier* dilambangkan dengan titik warna merah, sedangkan data bukan *outlier* dilambangkan dengan titik berwarna hitam. 9 data outlier tersebut terletak pada data Subulussalam, Aceh Besar, Aceh Barat, Bener Meriah, Bireuen, Aceh Jaya, Banda Aceh, Aceh Barat, Aceh Selatan.

3.3 Penentuan Jumlah *Cluster*

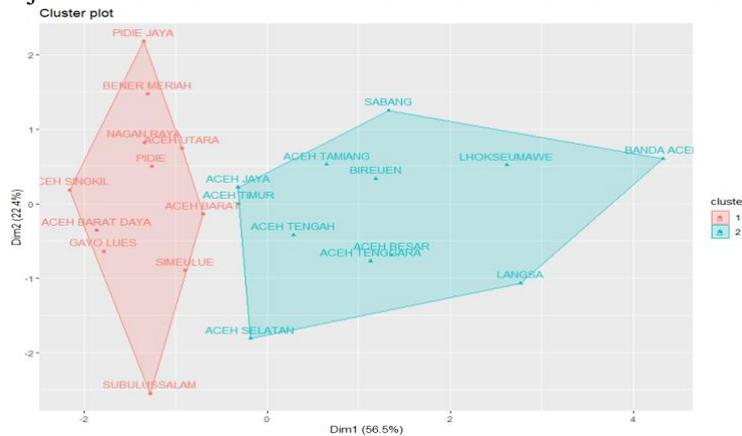


Gambar 3. Plot *Silhouette*

Menentukan jumlah *cluster* dengan metode yaitu *silhouette* untuk memperkirakan jumlah *cluster* terbaik. Dari *output* yang ada pada Gambar 3 terdapat saran dari metode *silhouette* melalui garis patah-patah yang menunjukkan pada angka 2 sebagai *cluster* optimum. Metode *silhouette* menggunakan nilai rata-rata jika semakin tinggi nilai maka semakin baik klasterisasi.

3.4 Analisis K-Medoids Clustering

Menggunakan analisis *cluster* dengan algoritma *K-Medoids* didapatkan hasil berupa plot *cluster* yang terdiri dari 2 *cluster*. Dalam pembentukan plot *cluster*, ukuran jarak yang digunakan yaitu ukuran jarak *euclidean*. Hal tersebut dikarenakan tipe ukuran jarak *euclidean* cocok untuk data yang tidak terjadi multikolinearitas atau ketika semua variabel tidak berkorelasi.



Gambar 4. Output K-Medoids Cluster

Dari *output* yang ada pada Gambar 4 dapat dilihat bahwa terdapat dua *cluster* yang didapatkan berdasarkan hasil perhitungan similaritas antar indikator kemiskinan di Provinsi Aceh. Pada plot hasil *cluster* terdapat 2 warna berbeda yaitu warna merah menunjukkan hasil *cluster* 1 dan warna biru menunjukkan hasil *cluster* 2. Setiap warna memiliki ciri khas tersendiri yang berbeda-beda.

Dari *output* yang ada pada Gambar 4 dapat dilihat juga anggota-anggota dalam setiap kelompok. Berikut merupakan anggota-anggota dalam setiap kelompok.

Tabel 4. Hasil Cluster K-Medoids

Cluster	Jumlah	Anggota Cluster
1	11	Aceh Singkil, Simeulele, Aceh Barat, Pidie, Gayo Lues, Aceh Utara, Bener Meriah, Pidie Jaya, Nagan Raya, Subulussalam, Aceh Barat Daya,
2	12	Aceh Selatan, Lhokseumawe, Aceh Tenggara, Langsa, Aceh Timur, Bireuen, Sabang, Aceh Tamiang, Aceh Tengah, Aceh Jaya, Banda Aceh, Aceh Besar

Dari *output* yang ada pada Table 4 dapat dilihat bahwa *cluster* 1 beranggotakan 11 kabupaten/kota, sedangkan *cluster* 2 beranggotakan 12 kabupaten/kota. Untuk mengetahui karakteristik dari setiap *cluster* akan dilakukan perhitungan rata-rata dari setiap variabel.

Selanjutnya, melakukan perhitungan nilai rata-rata dari masing-masing *cluster* untuk mengetahui karakteristik *cluster* berdasarkan faktor-faktor yang mempengaruhi kemiskinan. Terdapat empat kelompok dari seluruh Provinsi Aceh maka dapat diinterpretasikan sebagai berikut:

2. Dengan metode quan ditemukan adanya 9 data *outlier* dan 14 data *non-outlier*. 9 data outlier tersebut terletak pada data Subulussalam, Aceh Besar, Aceh Barat, Bener Meriah, Bireuen, Aceh Jaya, Banda Aceh, Aceh Barat, Aceh Selatan.
3. Dengan metode *silhouette* diperoleh jumlah *cluster* optimal sebanyak 2 *cluster*.
4. Berdasarkan pengelompokan analisis *K-Medoids clustering* diperoleh hasil yaitu *cluster 1* terdapat 11 kabupaten/kota, lalu *cluster 2* terdapat 12 kabupaten/kota.
5. Berdasarkan hasil profilisasi dan visualisasi peta dapat dilihat bahwa pada *cluster 1* terdapat 11 kabupaten/kota yang merupakan wilayah dengan karakteristik Persentase Penduduk Miskin dan IKK yang tinggi divisualisasikan dengan warna merah. Sedangkan *cluster 2* terdapat 12 kabupaten/kota yang merupakan wilayah dengan karakteristik Gini ratio, AHH, dan RLS yang tinggi divisualisasikan dengan warna biru.

6. UCAPAN TERIMAKASIH

Alhamdulillah, peneliti dapat menyelesaikan makalah penelitian mengenai "Implementasi *Clustering K-Medoids* Dalam Pengelompokan Kabupaten di Provinsi Aceh Berdasarkan Faktor yang Mempengaruhi Kemiskinan". Peneliti mengucapkan terimakasih kepada Dr. Edy Widodo, S.Si., M.Si. selaku Ketua Jurusan Statistika Universitas Islam Indonesia dan Dosen Pengampu beserta jajarannya, terima kasih atas limpahan ilmu yang diberikan dari awal hingga akhir serta bimbingan yang tiada hentinya. Freditasari Purwa Hidayat, M Dendi Alfitrah, Royhan Pina Putra sebagai tim peneliti. Serta dukungan dari pihak-pihak lain yang ikut menyemangati, memberikan saran, masukan serta dukungan dalam penyusunan penelitian ini.

DAFTAR PUSTAKA

- [1] Rustam, "Perencanaan Pertumbuhan Ekonomi Jawa Timur Dalam Rangka Mengurangi Angka Pengangguran Dan Kemiskinan", vol. 6, no. 1, 2010.
- [2] Bappenas, "Tujuan Pembangunan Berkelanjutan SDGs", <http://sdgs.bappenas.go.id/tujuan-1/>.
- [3] I. d. S. T. Raharjo, "Sustainable Development Goals (Sdgs) dan Pengentasan Kemiskinan", Jurnal Unpad, pp. 59.
- [4] E. Disty Wahyuli, "Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids," pp. 456-461, 2019.
- [5] H. Nugroho, "Kemiskinan, Ketimpangan dan Kesenjangan," Yogyakarta, Aditya media, 1995.
- [6] S. R. Ishartono, "Sustainable Development Goals (Sdgs) Dan Pengentasan Kemiskinan," Social Work Jurnal, vol. 6, no. 2, pp. 154 -272, 2020.
- [7] Y. N. Salafudin, Statistika Deskriptif, Pekalongan: STAIN Pekalongan Press, 2012.
- [8] R. R. Hocking, "Methods and Application of Linear Models," 1996.
- [9] T. Ryan, Modern Regression Methods, New York: John Wiley & Sons, 1997.
- [10] Supranto, "Analisis Multivariat Arti dan Interpretasi," Jakarta, PT Rineka Cipta, 2004.
- [11] D. T. Larose, Discovering Knowledge inData : An Introduction to Data Mining, Wiley-Interscience A John Wiley & Sons, Inc Publication, 2005.
- [12] H. Usman and N. Sobari, "Aplikasi Teknik Multivariat Untuk Riset Pemasaran," 2013.

- [13] S. Machfudhoh and N. Wahyuningsih, "Analisis Cluster Kabupaten/Kota Berdasarkan Pertumbuhan Ekonomi Jawa Timur," *Sains dan Seni Ponmits*, pp. 1-8, 2013.
- [14] Piptools, "Algoritma K-Medoids Clustering", <https://piptools.net/algoritma-k-medoids-clustering/>, 2015.
- [15] J. Han and M. Kamber, *Data Mining: Concept and Techniques*, Waltham: Morgan Kauffman Publisher, 2006.
- [16] A. W. Setyawati, "Implementasi Algoritma Partitioning Around Medoid (PAM) untuk Pengelompokan Sekolah Menengah Atas di DIY Berdasarkan Nilai Daya Serap Ujian Nasional," 2007.