

**Klasifikasi dan Klasterisasi Kemiskinan di Kabupaten Sabu Raijua Provinsi Nusa Tenggara Timur Menggunakan Machine Learning**

*Poverty Classification and Clusterization in Sabu Raijua District, East Nusa Tenggara Province Using Machine Learning*

Deddy Barnabas Lasfeto  
Politeknik Negeri Kupang  
email: deddylasfeto@gmail.com

---

**Abstract.** *The scope of this paper is focused on the socio-economy poverty problem in rural area. Household expenditure and income surveys provide data that are used for identifying and measuring the poverty status of households. The contribution of this work is the machine learning approach to assess and monitor the poverty status of households in the rural area of Indonesia. This approach takes into account all the household expenditure and income surveys that took place. This approach is accurate, inexpensive, and makes poverty identification cheaper and much closer to real-time. Data preprocessing and handling imbalanced data are major parts of this work. machine learning classification and clustering models are applied. The final machine learning classification and clustering model could transform efforts to track and target poverty. This work demonstrates how powerful and versatile machine learning can be, and hence, it promotes for adoption across many domains in both the private sector and government.*

*Keywords: society; poverty; data clusterization; data classification; machine learning.*

**Abstrak.** Ruang lingkup tulisan ini difokuskan pada masalah kemiskinan sosial ekonomi di pedesaan. Survei pengeluaran dan pendapatan rumah tangga menyediakan data yang digunakan untuk mengidentifikasi dan mengukur status kemiskinan rumah tangga. Kontribusi dari penelitian ini adalah pendekatan *machine learning* untuk memetakan dengan cara klasifikasi dan klasterisasi kemiskinan di pedesaan. Pendekatan ini memperhitungkan semua survei pengeluaran dan pendapatan rumah tangga yang dilakukan. Pendekatan machine learning lebih akurat, dan membuat identifikasi kemiskinan menjadi lebih mudah dan lebih mendekati real-time. Preprocessing data dan pengolahan data kemiskinan berbasis machine learning adalah bagian utama dari penelitian ini. Model klasifikasi dan klasterisasi berbasis machine learning terbukti dapat memberikan gambaran upaya untuk melihat kemiripan ciri kemiskinan. Hasil penelitian ini menunjukkan keunggulan machine learning dalam pemetaan kemiskinan desa di wilayah rural, oleh karena itu dapat diadopsi dalam banyak domain baik di sektor swasta maupun pemerintah.

Kata kunci: Masyarakat; kemiskinan; klasterisasi data; klasifikasi data; *machine learning*

---

## PENDAHULUAN

Salah satu tujuan pembangunan berkelanjutan (SDGs) adalah mengakhiri kemiskinan dalam segala bentuknya pada tahun 2030. Namun, penelitian terbaru Bank Dunia menunjukkan bahwa efek dari pandemi COVID-19 saat ini pasti akan muncul di sebagian besar negara hingga tahun 2030 [1]. Kemiskinan juga telah menjadi isu yang menyebar luas di Indonesia selama lebih dari satu dekade, dan masih menjadi salah satu isu yang paling banyak dibicarakan dan diperdebatkan hingga hari ini. Pemerintah terus mengupayakan berbagai cara untuk mengentaskan kemiskinan di seluruh tanah air.

Penanganan dan pengentasan kemiskinan ini juga merupakan suatu isu pokok pembangunan Kabupaten Sabu Raijua Provinsi Nusa Tenggara Timur. Kabupaten Sabu Raijua yang terdiri dari beberapa pulau termasuk Kabupaten Perbatasan dan/atau wilayah terluar berdasarkan Peraturan Presiden Republik Indonesia Nomor 179 Tahun 2014 tentang Rencana Tata Ruang Kawasan Perbatasan Negara di Provinsi Nusa Tenggara Timur. Persentase penduduk miskin di Kabupaten Sabu Raijua termasuk cukup tinggi yakni 30,18% pada tahun 2020 berada di atas rata-rata Provinsi sebesar 20,9% dan rata-rata Nasional sebesar 9,78% serta berada di posisi ke-21 dibandingkan dengan Kabupaten/Kota lain di provinsi NTT [2]. Sementara dalam dokumen Rencana Pembangunan Jangka Menengah Daerah (RPJMD) kabupaten Sabu Raijua periode tahun 2021-2026 ditargetkan persentase penduduk miskin akan menjadi 26,79% pada tahun 2026 [3]. Hal ini tentunya memerlukan pola penanganan yang lebih integrasi dengan menerapkan berbagai metode penelitian dan pengembangan untuk pengentasan masalah kemiskinan ini.

Melakukan survei rumah tangga adalah metode tradisional untuk memperkirakan kemiskinan suatu wilayah. Untuk mengidentifikasi rumah tangga miskin secara akurat, penilaian melalui survei dapat memakan waktu beberapa jam per rumah tangga. Secara umum model yang diterapkan adalah Proxy-means test (PMT) dimana model ini mengasumsikan bahwa terdapat hubungan linier antara variabel terikat dengan variabel bebas. Namun dalam kenyataannya asumsi ini tidak sepenuhnya terpenuhi, dimana bisa saja terjadi karakteristik rumah tangga yang disurvei tidak serta merta memiliki hubungan linier dengan tingkat kesejahteraan. Karakteristik rumah tangga dapat berinteraksi dengan cara yang kompleks, misalnya lantai tanah dapat menunjukkan kemiskinan bagi suatu wilayah tetapi tidak untuk wilayah yang lain. Oleh karena itu penelitian ini akan memanfaatkan metode machine learning yang merupakan bagian dari artificial intelligence untuk memetakan kemiskinan berdasarkan karakteristik yang dimiliki oleh rumah tangga tersebut. Data sosial ekonomi dianalisis menggunakan pendekatan machine learning dalam klasifikasi dan klusterisasi kemiskinan di wilayah kabupaten Sabu Raijua, agar memperoleh gambaran jelas kondisi kemiskinan, sehingga diharapkan semua program penanganan dan pengentasan kemiskinan akan lebih terarah dan tepat sasaran.

Artificial intelligence (AI) atau kecerdasan buatan adalah bidang ilmu komputer yang menekankan penciptaan mesin cerdas yang bekerja dan bereaksi seperti manusia. AI memungkinkan bagi mesin untuk belajar dari pengalaman, menyesuaikan dengan input baru dan melakukan tugas seperti manusia [4]. Dalam penelitian ini diterapkan Machine Learning yang merupakan cabang dari AI berdasarkan gagasan bahwa sistem dapat belajar dari data, mengidentifikasi pola, dan membuat keputusan dengan intervensi manusia yang manual. Dengan Machine Learning, komputer dapat menangani situasi baru melalui pelatihan mandiri, pengalaman, analisis, dan observasi [4]. Cara kerja machine learning adalah dengan mengumpulkan, memeriksa, dan membandingkan data berukuran kecil hingga besar untuk menemukan pola dan mengeksplorasi perbedaannya. Machine learning dibagi menjadi tiga jenis, yaitu supervised learning, unsupervised learning, dan reinforcement learning. Supervised learning adalah algoritma yang digunakan untuk memodelkan hubungan dan ketergantungan antara target output prediksi dan input sehingga kita dapat memprediksi nilai output untuk data baru berdasarkan hubungan dan ketergantungan yang telah terbentuk sebelumnya. Unsupervised learning adalah algoritma machine learning yang memiliki fungsi utama untuk mendeteksi pola dan pemodelan deskriptif. Algoritma ini tidak memiliki kategori atau label output pada data (tidak memiliki data latih dan data uji). Sedangkan reinforcement learning adalah algoritma yang bertujuan untuk mengambil tindakan yang akan memaksimalkan output dan meminimalkan resiko dengan cara mengamati dan mengumpulkan interaksi dengan environmentnya [4].

Revolusi digital terus menghasilkan banyak data, yang memberikan peluang baru untuk menangkap informasi tentang kondisi sosial ekonomi di berbagai daerah untuk menyimpulkan kemajuan pembangunan [5]. Ketersediaan data tentang kemiskinan dan ketimpangan ekonomi di Kabupaten Sabu Raijua masih terbatas, sementara data tersebut perlu digunakan untuk memantau perubahan tingkat kesejahteraan, serta untuk mengukur dampak program pemerintah. Penelitian yang akan dilakukan ini untuk mengaplikasikan sistem kecerdasan buatan dengan memanfaatkan machine learning untuk memetakan kemiskinan menggunakan kumpulan data terkait yakni citra satelit dan data sosial ekonomi. Hasil pengukuran kemiskinan dengan artificial intelligence ini dimaksudkan untuk melengkapi data Survei Sosial Ekonomi Nasional yang telah dikumpulkan Badan Pusat Statistik (BPS).

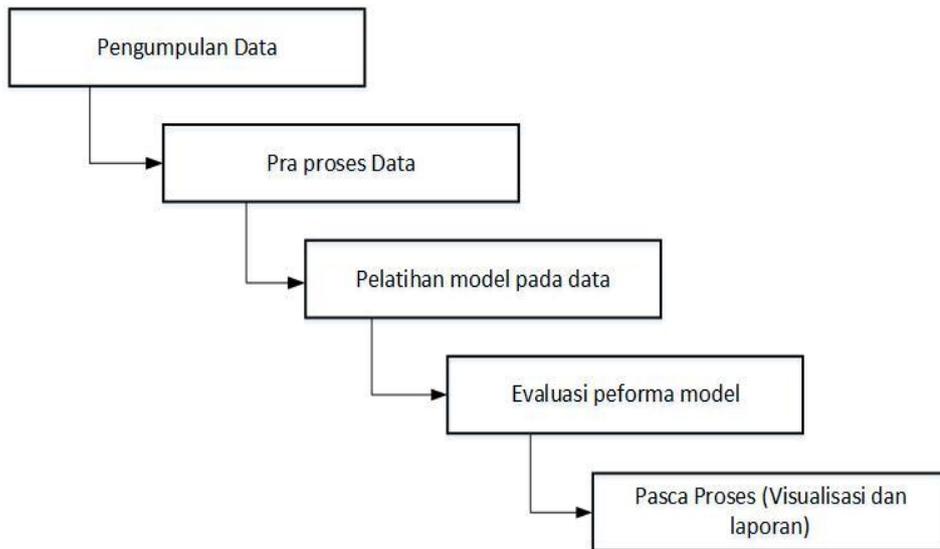
Banyak penelitian telah membuktikan machine learning dapat digunakan untuk pemetaan kemiskinan. Pengusul mengelompokkan studi pustaka tentang penerapan machine learning untuk analisis kemiskinan ini berdasarkan sumber datanya. Berdasarkan hasil sensus: Penelitian untuk mengklasifikasikan level kemiskinan dari data sosial ekonomi [5, 6, 7, 8, 9], prediksi level kemiskinan [10,11]. Berdasarkan analisis citra satelit: penelitian lain menerapkan machine learning dengan tujuan mengidentifikasi fitur dari citra satelit yang dapat menggambarkan sekitar 75% variasi status ekonomi tingkat lokal [12]. Peta kemiskinan juga dihasilkan dari citra satelit ini, yang penting untuk penargetan kemiskinan dan penyediaan layanan publik, khususnya di negara-negara berpenghasilan rendah [1, 13]. Demikian pula, data citra tentang kondisi infrastruktur wilayah digunakan untuk memprediksi indikator mata pencaharian utama telah memberikan hasil yang baik [14]. Masalah klasifikasi kemiskinan berdasarkan data citra satelit telah dianalisis pula [12], demikian pula dengan pemetaan kemiskinan [1, 13, 14], dan evaluasi program [15].

Secara nasional indikator umum pengukur kemiskinan adalah kemiskinan makanan atau pangan (GKM) dan kemiskinan non pangan (GKNM) meliputi perumahan, sandang, pendidikan dan kesehatan. Selain itu untuk melihat kemajuan pembangunan suatu wilayah perlu pula dilihat infrastruktur umum seperti sarana penerangan, pendidikan, dan transportasi [2]. Berbagai penelitian terdahulu di atas melakukan analisis kemiskinan berdasarkan pemodelan yang parsial dengan menggunakan data sekunder sosial ekonomi, atau menggunakan data citra satelit saja. Penggunaan metode secara parsial ini bila dikaitkan dengan penanganan dan pengentasan kemiskinan secara nasional tentunya menghasilkan informasi yang tidak komplit, misalnya dengan menggunakan data sosial ekonomi, maka informasi yang bisa dihasilkan adalah informasi kemiskinan makanan (GKM) sedangkan kemiskinan non makanan (GKNM) yang dihasilkan akan sulit dipetakan. Masalah kondisi rumah tinggal seperti yang telah digambarkan sebelumnya pada suatu wilayah bisa berbeda persepsinya terhadap kemiskinan meskipun memiliki model struktur bangunan yang sama, hal ini dipengaruhi oleh persepsi sosial dan budaya daerah yang berbeda [8].

Penelitian yang diusulkan ini melihat dari Perbedaan-perbedaan analisis kemiskinan yang dipublikasikan di atas sehingga akan dikembangkan model pemetaan berbasis machine learning menggunakan data sosial ekonomi, yang diharapkan menjadi suatu solusi yang lebih tepat bagi penanganan pemetaan kemiskinan di suatu wilayah.

## **METODOLOGI**

Langkah-langkah proses penelitian yang dilakukan ditunjukkan pada Gambar 1 berikut ini.



Gambar 1. Tahapan-tahapan penelitian

Pengambilan data survei sosial ekonomi berbasis pada indikator kemiskinan secara nasional dari Badan Pusat Statistik sebanyak 14 indikator dimana Pemutakhiran data PSE05 dalam PPLS 2008 menggunakan pendekatan karakteristik rumah tangga dengan 14 variabel kualitatif penjas kemiskinan, yaitu 1. luas lantai per kapita, 2. jenis lantai, 3. jenis dinding, 4. fasilitas buang air besar, 5. sumber air minum, 6. sumber penerangan, 7. bahan bakar, 8. pembelian daging/ayam/susu, 9. frekuensi makan, 10. pembelian pakaian baru, 11. kemampuan berobat, 12. lapangan usaha kepala rumah tangga, 13. pendidikan kepala rumah tangga, dan 14. aset yang dimiliki.

Indikasi kemiskinan adalah sebagai berikut:

1. Luas lantai bangunan tempat tinggal kurang dari 8 m<sup>2</sup> per orang
2. Jenis lantai tempat tinggal terbuat dari tanah/bambu/kayu murahan
3. Jenis dinding tempat tinggal dari bambu/ rumbia/ kayu berkualitas rendah/tembok tanpa diplester.
4. Tidak memiliki fasilitas buang air besar/ bersama-sama dengan rumah tangga lain.
5. Sumber penerangan rumah tangga tidak menggunakan listrik.
6. Sumber air minum berasal dari sumur/ mata air tidak terlindung/ sungai/ air hujan.
7. Bahan bakar untuk memasak sehari-hari adalah kayu bakar/ arang/ minyak tanah
8. Hanya mengkonsumsi daging/ susu/ ayam dalam satu kali seminggu.
9. Hanya membeli satu stel pakaian baru dalam setahun
10. Hanya sanggup makan sebanyak satu/ dua kali dalam sehari
11. Tidak sanggup membayar biaya pengobatan di puskesmas/ poliklinik
12. Sumber penghasilan kepala rumah tangga adalah: petani dengan luas lahan 500 m<sup>2</sup>, buruh tani, nelayan, buruh bangunan, buruh perkebunan dan atau pekerjaan lainnya dengan pendapatan dibawah Rp. 600.000,- per bulan
13. Pendidikan tertinggi kepala rumah tangga: tidak sekolah/ tidak tamat SD/ tamat SD.
14. Tidak memiliki tabungan/ barang yang mudah dijual dengan minimal Rp. 500.000,- seperti sepeda motor kredit/ non kredit, emas, ternak, kapal motor, atau barang modal lainnya.

Sampel penelitian ini adalah Kecamatan di Kabupaten Sabu Raijua yang memiliki distribusi luas wilayah paling besar dan jumlah desa terbanyak, serta jumlah penduduk terbanyak untuk dijadikan sampel pada penelitian ini yakni Kecamatan Sabu Barat, dengan data yang ditunjukkan dalam tabel 1 dan tabel 2 berikut ini.

Tabel 1. Pembagian Wilayah Administrasi di Kabupaten Sabu Raijua

No	Kecamatan	Ibukota	Jumlah Desa	Jumlah Kelurahan	Luas Wilayah (Km <sup>2</sup> )	Distribusi Luas Wilayah
1	Raijua	Ledeunu	3	2	39,05	8 %
2	Sabu Barat	Seba	17	1	185,16	40 %
3	Hawu Mehara	Tanajawa	10	-	62,81	14 %
4	Sabu Timur	Bolow	8	2	37,21	8 %
5	Sabu Liae	Eilogo	12	-	57,62	13 %
6	Sabu Tengah	Eimmadeke	8	-	78,62	17 %
<b>Jumlah</b>			<b>58</b>	<b>5</b>	<b>460,47</b>	<b>100 %</b>

Sumber : BPS Kabupaten Sabu Raijua 2020

Tabel 2. Proyeksi Penduduk Dirinci Menurut Jenis Kelamin

No	Kecamatan	Laki-Laki	Perempuan	Jumlah	Rasio Jenis Kelamin	Kepadatan Penduduk Per Km <sup>2</sup>
1	Raijua	4,499	4,451	8,950	101.1	235
2	Sabu Barat	17,056	16,169	33,225	105.5	179
3	Hawu Mehara	9,561	9,103	18,664	105.0	297
4	Sabu Timur	4,553	4,384	8,937	103.9	240
5	Sabu Liae	5,394	5,324	10,718	101.3	186
6	Sabu Tengah	4,549	4,284	8,833	106.2	112
<b>Jumlah</b>		<b>45,612</b>	<b>43,715</b>	<b>89,327</b>	<b>103.8</b>	<b>208.2</b>

Sumber: BPS, "Buku Kabupaten Sabu Raijua Dalam Angka 2021" (diolah)

Data kuantitatif kemiskinan yang diperoleh selanjutnya dianalisis menggunakan metode machine learning yakni yakni klasifikasi kemiskinan menggunakan algoritma K\_nearest Neighbor (K-NN) dan klusterisasi kemiskinan menggunakan algoritma K-means clustering.

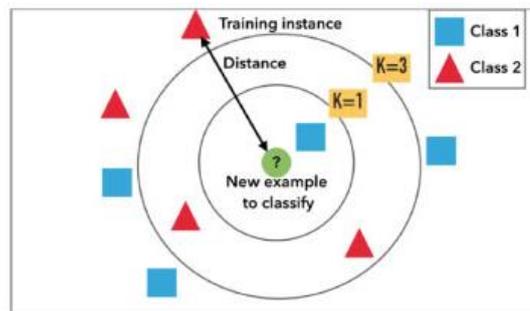
### Metode klasifikasi K-Nearest Neighbor (K-NN) dalam Machine Learning

K-Nearest Neighbors adalah salah satu algoritma machine learning paling dasar supervised learning, namun sangat penting. Algoritma machine learning supervised learning adalah salah satu jenis algoritma machine learning yang bergantung pada data masukan berlabel untuk mempelajari fungsi yang mampu menghasilkan keluaran setiap saat ketika data tak berlabel baru diberikan sebagai masukan. K-NN banyak digunakan karena bersifat non-parametrik yang berarti tidak membuat asumsi mendasar tentang distribusi data. K-NN menggunakan seluruh dataset dalam fase pelatihannya. Setiap kali prediksi diperlukan untuk data baru, K-NN menelusuri seluruh himpunan data pelatihan untuk melihat data yang paling mirip dengan k dan data yang paling mirip akhirnya dikembalikan sebagai prediksi. contoh sederhana, jika gambar buah apel terlihat lebih mirip dengan gambar buah pir dan ceri (buah-

buah) daripada gambar monyet, kucing atau tikus (hewan), maka kemungkinan besar apel adalah buah.

Tujuan utama algoritma K-Nearest Neighbor adalah untuk mengklasifikasikan titik data baru dengan membandingkannya dengan semua titik data yang terlihat sebelumnya. Klasifikasi K kasus sebelumnya yang paling mirip digunakan untuk memprediksi klasifikasi titik data saat ini. Dalam algoritma K-NN, kualitas prediksi sepenuhnya tergantung pada ukuran jarak. Dengan demikian, algoritma ini cocok untuk aplikasi dengan basis domain pengetahuan yang memadai sehingga dapat membantu dalam memilih ukuran yang sesuai. K-NN dapat digunakan untuk masalah prediksi regresi dan klasifikasi. Namun, dalam prakteknya, algoritma ini banyak digunakan dalam masalah klasifikasi. Algoritma ini paling sering digunakan untuk kemudahan interpretasi dan waktu kalkulasi yang lebih singkat.

Algoritma K-NN bekerja atas dasar kesamaan fitur. Klasifikasi titik data tertentu ditentukan oleh seberapa mirip fitur di luar sampel yang mirip dengan set pelatihan.



Gambar 2. klasifikasi K-NN

Gambar di atas menunjukkan model klasifikasi K-NN. Jika mempertimbangkan tetangga terdekat dari sampel uji (?), maka yang terdekat adalah kotak biru (Kelas 1) atau  $k = 1$  yang berada di dalam lingkaran dalam. Jika menganggap  $k = 3$ , maka akan melihat 2 segitiga merah dan hanya 1 kotak biru yang berada di bawah lingkaran luar. Dengan demikian, sampel uji sekarang diklasifikasikan sebagai segitiga merah (Kelas 2). Demikian pula, jika menggunakan  $k = 5$ , kelas ini ditetapkan ke kelas pertama (3 kotak vs. 2 segitiga di luar lingkaran luar).

Sebuah kasus dapat diklasifikasikan dengan frekuensi mayoritas dari tetangganya. Kasus tersebut kemudian didistribusikan ke kelas yang paling umum di antara K tetangga terdekatnya yang diukur dengan fungsi jarak. Ukuran jarak yang digunakan untuk mengetahui tetangga terdekat adalah menggunakan salah satu rumus berikut:

$$\begin{array}{l}
 \text{Euclidean} \quad \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad \text{Manhattan} \quad \sum_{i=1}^k |x_i - y_i| \quad \text{Minkowski} \quad \left( \sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q}
 \end{array}$$

Dalam penelitian ini, digunakan Rumus Euclidean adalah sebagai berikut:

$$D = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Dengan memeriksa data, dapat memilih nilai optimal terbaik untuk K. Umumnya, nilai K yang besar lebih akurat karena cenderung mengurangi noise atau tidak valid secara keseluruhan. Dalam penelitian ini digunakan nilai  $K = 3$ .

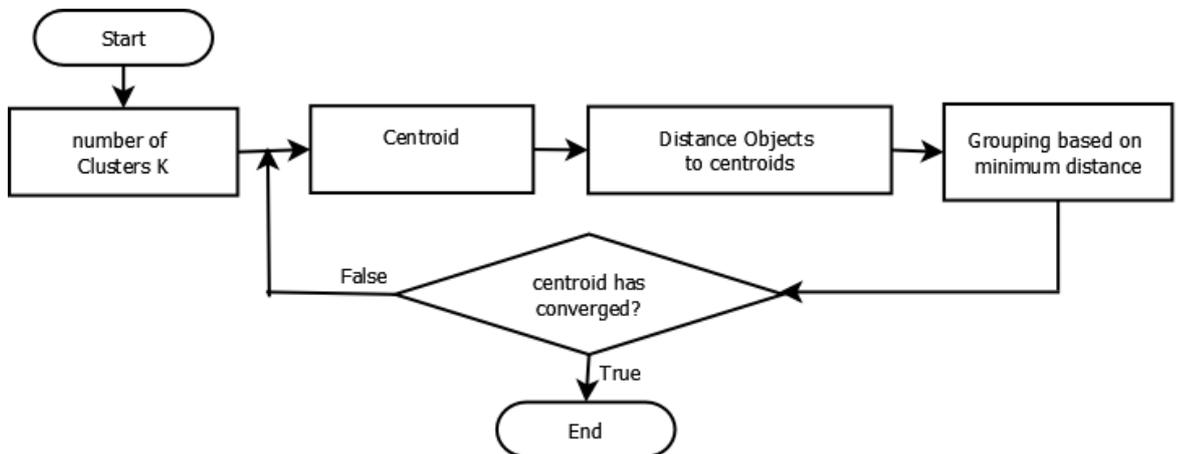
### Metode klasterisasi dengan K-Means clustering

Algoritma K-means membagi clustering berbasis jarak yang membagi data ke dalam sejumlah cluster. Metode ini hanya bekerja pada atribut numerik, dan parameter Jarak digunakan untuk mencari tingkat kemiripan.

Langkah-langkah dalam algoritma K-means clustering adalah sebagai berikut:

1. Tentukan jumlah k kluster
2. Jumlah K kluster ini ditentukan secara bebas tetapi harus memenuhi sejumlah syarat dimana salah satu syarat yang umum adalah nilai k harus lebih kecil dari jumlah data;
3. Pilih titik secara acak sebanyak k buah dimana titik ini menjadi pusat (centroid) dari masing-masing kelompok (cluster);
4. Hitung jarak dan alokasikan masing-masing data ke centroid (rata-rata terdekat);
5. Tentukan centroid baru / rata-rata dari data yang ada dalam masing-masing cluster;
6. Kembali mengulangi langkah ke-3 apabila masih ada data yang berpindah cluster atau ada perubahan nilai centroid. Jika tidak ada perubahan maka hentikan proses clustering;

Langkah-langkah ini dapat dilihat ada flow-chart berikut ini.



Gambar 3. Flow chart algoritma K-means clustering

Penentuan jarak terdekat antara data dengan centroid pada umumnya menggunakan euclidean distance:

$$D(a,b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2}$$

dimana: D adalah jarak, n adalah jumlah dimensi (atribut),  $a_k$  dan  $b_k$  adalah atribut ke-k dari objek data p dan q

### HASIL DAN PEMBAHASAN

Survei indikator kemiskinan ini mengambil sampel 500 Kepala Keluarga yang tersebar di 17 Desa yang ada di Kecamatan Sabu Barat. Pengambilan data dilaksanakan pada bulan September 2022. Data responden ini direkap untuk setiap indikator kemiskinan, dan dikalkulasi secara total setiap responden (kepala Keluarga) yang memenuhi kriteria kemiskinan pada masing-masing indikator untuk seluruh Desa yang ada di kecamatan Sabu Barat. Tabel di bawah ini menunjukkan rekapan hasil survei tersebut.

Tabel 3. Rekapitan Hasil Survei

No.	Nama Desa	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14
1	Raeloro	26	16	27	8	2	10	28	28	27	27	27	27	6	27
2	Raenyale	0	21	28	2	0	0	28	28	23	20	28	25	8	12
3	Ledeana	5	25	26	5	0	0	27	21	16	20	27	22	15	20
4	Titinalede	22	22	22	6	5	5	22	19	10	20	22	20	12	10
5	Djadu	23	18	23	5	18	17	23	22	15	10	23	23	13	13
6	Menia	5	28	36	9	11	5	36	36	34	23	36	35	10	33
7	Ledekepaka	9	27	29	13	0	0	28	15	14	22	29	26	14	24
8	Depe	29	29	29	9	13	28	29	28	13	29	29	29	7	28
9	Roboaba	3	18	26	2	0	26	26	24	19	15	26	20	3	20
10	Teriwu	1	21	22	0	4	21	22	22	14	17	22	20	12	22
11	Raemude	28	28	28	18	17	24	28	27	28	25	28	26	20	28
12	Raekore	27	28	28	7	15	28	28	28	28	21	27	28	13	28
13	Raedewa	14	22	31	1	3	1	35	35	26	32	35	32	11	12
14	Nadawawi	19	25	27	7	4	12	26	31	27	22	21	28	5	11
15	Mebba	1	9	27	7	1	12	22	27	8	19	28	22	12	15
16	Delo	28	28	28	4	10	26	28	28	28	5	26	28	15	28
17	Raemadia	23	30	31	4	5	1	30	24	28	19	30	30	10	15

Sumber: Data primer hasil penelitian (diolah)

Dari tabel 3 dapat dijelaskan bahwa lebih dari 90% seluruh Desa di Kecamatan Sabu Barat memiliki data Kepala Keluarga yang memenuhi setiap indikator kemiskinan. Sebagai contoh, pada Desa Raeloro terdapat 26 Kepala Keluarga yang memiliki rumah dengan luas lantai bangunan tempat tinggal kurang dari 8 m<sup>2</sup> per kapita (I1); sebanyak 16 Kepala Keluarga memiliki rumah dengan Jenis lantai tempat tinggal terbuat dari tanah/bambu/kayu murahan (I2); Sebanyak 27 Kepala Keluarga yang Hanya sanggup makan sebanyak satu/ dua kali dalam sehari (I10); Data detail setiap desa dapat dilihat pada tabel di atas.

Dalam memetakan kemiskinan sebagai tujuan penelitian ini, maka dilakukan analisis berbasis artificial intelligence yakni pendekatan machine learning untuk melakukan klasifikasi dan klusterisasi indikator-indikator kemiskinan pada semua desa di Kecamatan Sabu Barat ini. Hasil analisis klasifikasi dan klusterisasi akan diuraikan berikut ini.

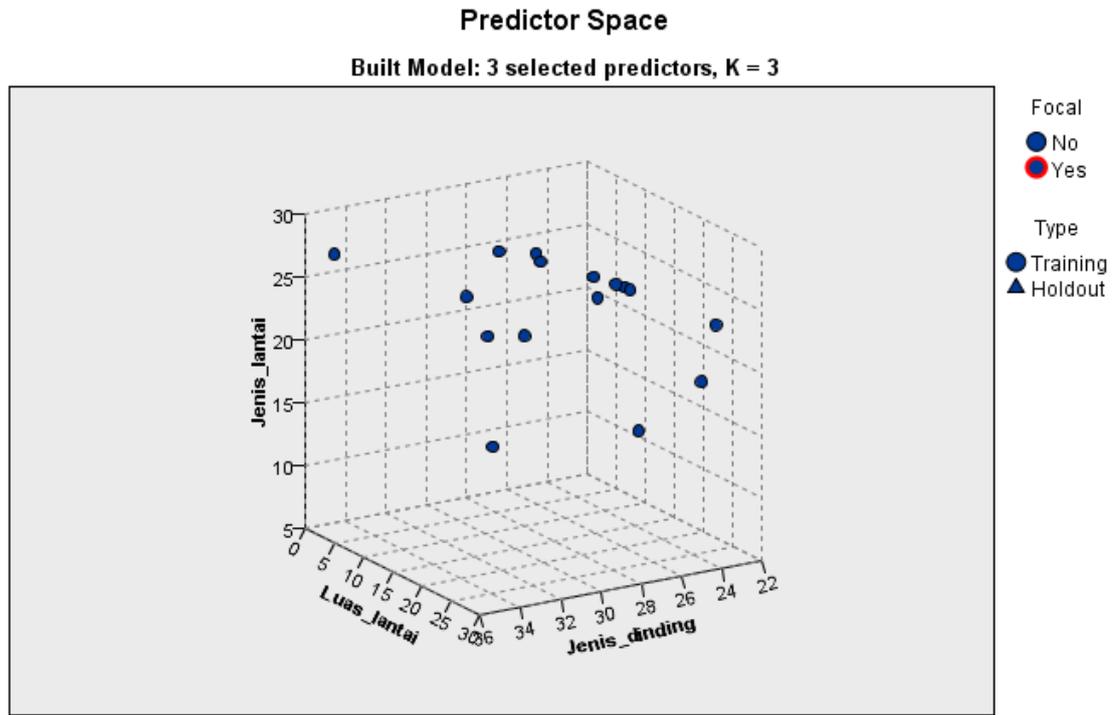
### Klasifikasi kemiskinan

Proses analisis klasifikasi kemiskinan pada masing-masing desa menggunakan teknik machine learning dengan Algoritma K-Nearest Neighbor. Secara ringkas, hasil analisis klasifikasi dengan algoritma K-Nearest Neighbor adalah sebagai berikut:

Tabel 4. Rekapitan data yang diproses

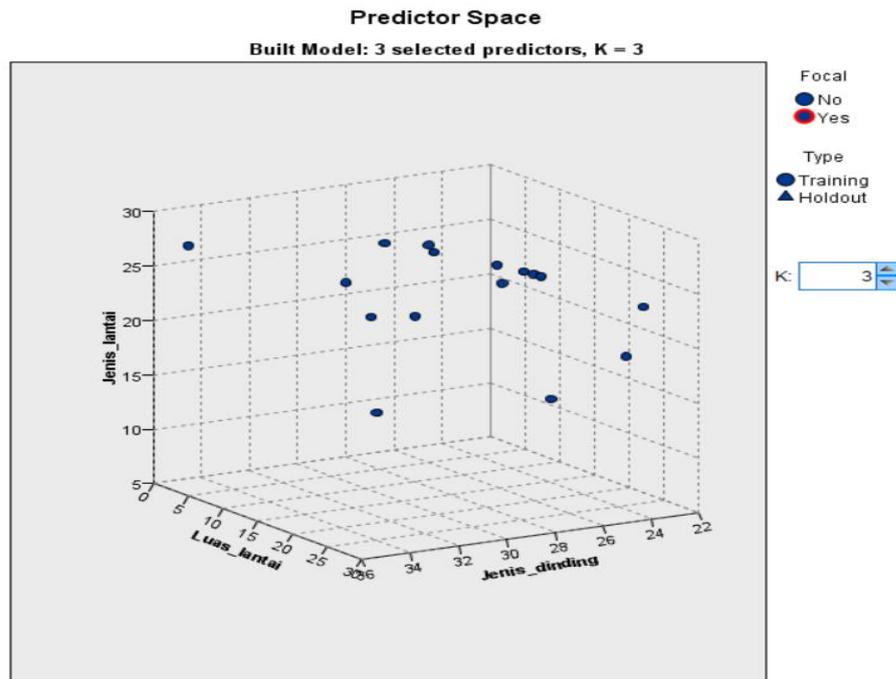
	N	Percent
Sample		
Training	16	94.1%
Holdout	1	5.9%
Valid	17	100.0%
Excluded	0	
Total	17	

Terlihat semua data sebanyak 17 data dilatih. Grafik prediktor dari analisis klasifikasi K-nearest neighbor ini ditunjukkan dalam grafik berikut:



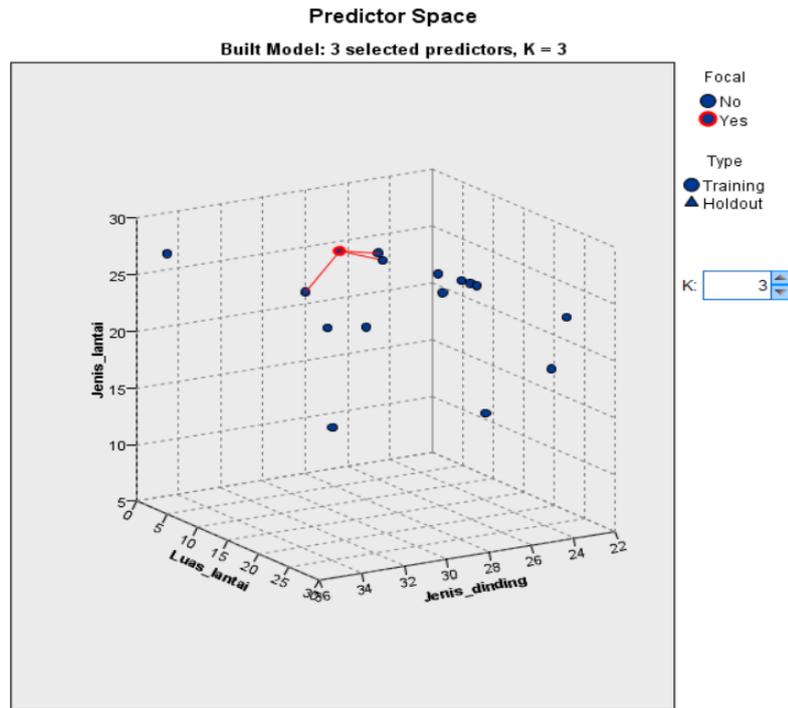
Gambar 4. Grafik predictor klasifikasi

Hasil klasifikasi ditunjukkan pada grafik berikut:



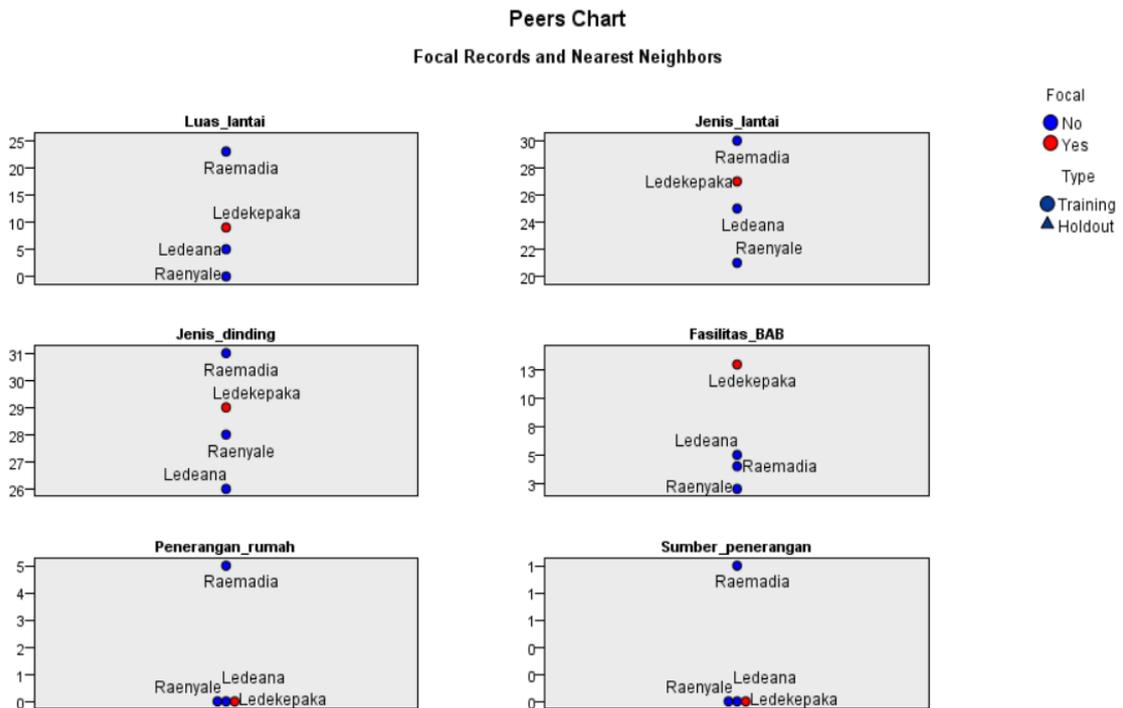
Gambar 5. Hasil klasifikasi

Salah satu data training untuk melihat kedekatannya dengan tiga data sebagai tetangga terdekat (nearest neighbour) berdasarkan analisis:



Gambar 6. Hasil analisis salah satu desa

Pada gambar di atas, diambil Desa Ledekkepaka untuk mencari kemiripan karakteristik kemiskinan dengan tiag desa lainnya pada setiap indicator kemiskinan.



Gambar 7. Pemetaan klasifikasi

Dari Grafik ini, data yang dipilih adalah Desa Ledekkepaka, pada Inidikator Luas lantai (I1), terlihat bahwa Desa Ledekkepaka memiliki kedekatan atau kemiripan dengan Desa Raemadia, Desa Ledeana, dan Desa Raenyale.

Pada Indikator Fasilitas buang air besar (I4), Desa Ledekepaka memiliki kemiripan dengan Desa Ledeana, Desa Raemadia, dan Desa raenyale. Secara keseluruhan, tabel jarak setiap indicator kemiskinan dari Desa Ledekepaka ini dengan tetangga terdekatnya ditunjukkan dalam tabel di bawah ini.

Tabel 5. Hasil klafisikasi dengan algoritma KNN

Displayed for Initial Focal Records						
Focal Record	Nearest Neighbors			Nearest Distances		
	1	2	3	1	2	3
Ledekepaka	Ledeana	Raemadia	Raenyale	1.407	2.366	2.428

Dari hasil analisis ini terlihat bahwa Desa Ledekepaka memiliki kesamaan ciri kemiskinan dengan tiga desa lainnya yakni desa Ledeana, desa Raemadia, dan desa Raenyale. Akan tetapi, yang paling dekat atau dikatakan hampir sama dengan Desa Ledekepaka adalah Desa Ledeana, terlihat dari jarak data keduanya adalah paling dekat yakni 1,407.

### Klasterisasi kemiskinan

Klasterisasi diperlukan untuk mengelompokkan Desa - desa yang memiliki kesamaan ciri kemiskinan sehingga dapat dijadikan rujukan pihak terkait dalam mengambil Langkah-langkah kebijakan penanganan kemiskinan di Kabupaten sabu Raijua.

Analisi klaster ini menggunakan algoritma K-Means Clustering. Dari analisis data dengan algoritma ini diperoleh hasil pada tabel 6. Tabel 6 menunjukkan cluster awal (initial cluster) dan centroid dari masing-masing cluster untuk setiap variabel. Langkah-langkah iterasi yang ditempuh sistem K-means clustering ditunjukkan dalam tabel 7. Dari tabel 7, diketahui bahwa pengolahan clustering ini, menempuh 2 kali iterasi, dan sesudah itu data mengalami konvergen (tidak berubah), dan diketahui bahwa jarak minimumnya adalah 6,437. Oleh karena sudah konvergen, maka kita bisa lihat pada cluster akhir dari analisis K-means clustering (tabel 8). Tabel 8 menunjukkan cluster akhir dimana sudah tidak ada lagi perubahan centroid dari masing-masing cluster. Grafik dari final cluster ini ditunjukkan pada gambar 8. Dari analisis cluster ini dapat diketahui bahwa Klaster 1 dengan tingkat kemiskinan yang relative lebih tinggi, klaster 2 relative sedang, dan klaster 3 dengan tingkat kemiskinan relative rendah.

Dari hasil analisis klasterisasi ini, dapat diketahui bahwa pada klaster 1 beberapa indikator kemiskinan yang sangat signifikan berpengaruh adalah: (1) variable Jenis dinding rumah yakni mayoritas dinding rumah masyarakat menggunakan bambu/rumbia/kayu berkualitas rendah atau tembok tanpa diplester; (2) variable bahan bakar yang dikonsumsi yakni bahan bakar yang digunakan untuk memasak sehari-hari adalah kayu bakar. Variabel ini dapat didiskusikan lebih lanjut pada penelitian-penelitian selanjutnya karena pengaruh variable ini terhadap kemiskinan perlu dilihat dari kondisi geografis dan budaya setempat; (3) variable biaya pengobatan yakni mayoritas masyarakat di klaster 1 tidak sanggup membayar biaya pengobatan di Puskesmas/Klinik/Rumah sakit. Variabel ini sudah diintervensi oleh pemerintah lewat program janiman Kesehatan yakni Kartu Indonesia Sehat; (4) variable pendapatan Kepala keluarga dimana mayoritas Kepala Keluarga berpendapatan bulanan lebih kecil dari Rp.600.000,-. Variabel ini dapat diintervensi pemerintah melalui program-program pemberdayaan dan peningkatan ekonomi keluarga.

Tabel 6. Initial cluster centers

	Cluster		
	1	2	3
Zscore(Luas_lantai)	-.13199	1.12454	-1.29877
Zscore(Jenis_lantai)	-.21815	.84142	-2.51387
Zscore(Jenis_dinding)	1.01184	.13720	-.15435
Zscore(Fasilitas_BAB)	-1.18467	2.61944	.15796
Zscore (Penerangan_rumah)	-.53103	1.68624	-.84778
Zscore (Sumber_penerangan)	-1.06817	1.03060	-.06441
Zscore (Bahan_bakar_konsumsi )	1.90407	.14760	-1.35794
Zscore (Konsumsi_daging_susu )	1.66223	.17497	.17497
Zscore (Membeli_pakaian_per_tahun)	.63042	.88558	-1.66610
Zscore (Frekuensi_makan_sehari)	1.79201	.71499	-.20816
Zscore (Biaya_pengobatan)	1.88450	.17263	.17263
Zscore(Penghasilan_KK)	1.38784	.01347	-.90277
Zscore(Pendidikan_KK)	.01394	2.14662	.25090
Zscore (Tabungan_barang)	-1.18521	1.02285	-.77120

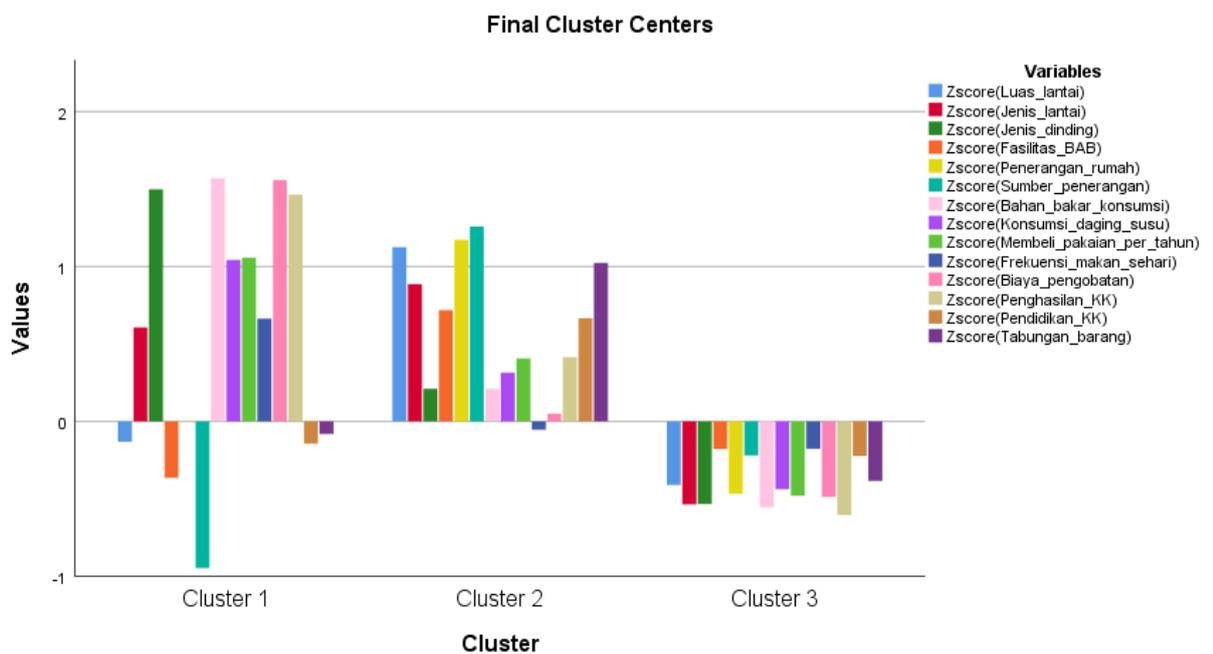
Tabel 7. Iterasi klusterisasi

Iteration	Change in Cluster Centers		
	1	2	3
1	2.278	2.675	2.908
2	.000	.000	.000

- a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 6.437.

Tabel 8. Final cluster centers

	Cluster		
	1	2	3
Zscore(Luas_lantai)	-.13199	1.12454	-.41022
Zscore(Jenis_lantai)	.60596	.88557	-.53602
Zscore(Jenis_dinding)	1.49775	.21009	-.53336
Zscore(Fasilitas_BAB)	-.36418	.71738	-.17770
Zscore (Penerangan_rumah)	-.00311	1.17152	-.46768
Zscore (Sumber_penerangan)	-.94650	1.25872	-.21954
Zscore (Bahan_bakar_konsumsi )	1.56951	.21033	-.55499
Zscore (Konsumsi_daging_susu )	1.04254	.31440	-.43852
Zscore (Membeli_pakaian_per_tahun)	1.05570	.40714	-.47957
Zscore (Frekuensi_makan_sehari)	.66371	-.05430	-.17739
Zscore (Biaya_pengobatan)	1.55843	.05035	-.48767
Zscore(Penghasilan_KK)	1.46420	.41433	-.60499
Zscore(Pendidikan_KK)	-.14404	.66559	-.22302
Zscore (Tabungan_barang)	-.08118	1.02285	-.38479



Gambar 8. Hasil klasterisasi

Hasil pengelompokan (klustersiasi) semua desa ditampilkan dalam tabel berikut ini.

Tabel 9. Klasterisasi Desa untuk semua indicator kemiskinan

Number of Cases in each Cluster		
Cluster	1	3.000
	2	4.000
	3	10.000
Valid		17.000
Missing		.000

Dari hasil analisis ini diperoleh bahwa yang termasuk dalam Cluster 1 sebanyak 3 desa yakni: Desa Menia, Desa Raedewa; Desa Raemadia. Dengan demikian diperoleh bahwa ketiga desa ini memiliki kesamaan ciri kemiskinan sehingga bisa menjadi rujukan pengentasan kemiskinan pada ketiga desa ini dengan pola yang sama.

Yang termasuk dalam Cluster 2 sebanyak 4 desa yakni: Desa Depe; Desa Raemude; Desa Raekore; dan Desa Delo dimana keempat Desa ini memiliki ciri kemiskinan yang sama. Yang termasuk dalam Cluster 3 sebanyak 10 Desa yakni: Desa Raeloro; Desa Raenyale; Desa Ledean; Desa Titinalede; Desa Djadu; Desa Ledekpeka; Desa Roboaba; Desa Teriwu; Desa Nadawawi; dan Desa Meba.

## KESIMPULAN

Dengan menggunakan algoritma klasifikasi dan klasterisasi, dapat dipetakan kondisi kesamaan atau kemiripan ciri kemiskinan dari setiap desa yang ada di Kecamatan Sabu Barat Kabupaten Sabu Raijua. Kondisi kemiskinan di desa dapat dianalisis dan prioritas dalam pelaksanaan program sosial ekonomi di setiap cluster desa direkomendasikan dalam penelitian ini sesuai dengan empat belas (14) indikator sosial ekonomi untuk program pengentasan kemiskinan dan mengembangkan strategi yang komprehensif bagi masyarakat.

Analisis klasifikasi dan klasterisasi machine learning digunakan sebagai alat pendukung keputusan untuk memperoleh pemahaman data kemiskinan setiap desa dan menggambarkan profil kluster kemiskinan di masyarakat dalam hal indikator sosial ekonomi signifikan yang ada dalam data. Kluster desa berdasarkan analisis indikator-indikator kemiskinan yang ada dikelompokkan menjadi tiga (3) kluster yang digambarkan sebagai: kluster kemiskinan tinggi, kluster kemiskinan sedang, dan kluster kemiskinan rendah. Analisis ini dapat digunakan sebagai dasar dalam merekomendasikan program yang akan mengatasi kondisi kemiskinan yang paling mendesak di setiap kluster.

## UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Direktorat Akademik Pendidikan Tinggi Vokasi Kementerian Pendidikan Kebudayaan Riset dan Teknologi yang mendanai penelitian ini.

## DAFTAR PUSTAKA

- Asian Development Bank. (2021). Mapping the Spatial Distribution of Poverty Using Satellite Imagery in Thailand (Issue April).
- Badan Pusat Statistik. (2021). Data dan Informasi Kemiskinan Kabupaten/Kota tahun 2021.

- Pemerintah Kabupaten Sabu Raijua. (2021). Rencana Pembangunan Jangka Menengah Daerah tahun 2021 - 2026.
- Hahn, P. (2019). Artificial intelligence and machine learning. *Handchirurgie Mikrochirurgie Plastische Chirurgie*, 51(1), 62-67. <https://doi.org/10.1055/a-0826-4789>
- Alsharkawi, A., Al-Fetyani, M., Dawas, M., Saadeh, H., & Alyaman, M. (2021). Poverty classification using machine learning: The case of Jordan. *Sustainability (Switzerland)*, 13(3), 1-16. <https://doi.org/10.3390/su13031412>
- Effendy, F., & Purbandini, P. (2018). Klasifikasi Rumah Tangga Miskin Menggunakan Ordinal Class Classifier. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 4(1), 30-36. <https://doi.org/10.25077/teknosi.v4i1.2018.30-36>
- Nofriani, N. (2020). Machine Learning Application for Classification Prediction of Household's Welfare Status. *JITCE (Journal of Information Technology and Computer Engineering)*, 4(02), 72-82. <https://doi.org/10.25077/jitce.4.02.72-82.2020>
- Poreddy, D., Reddy, E. V. V., Prasad, S. V., & Reddy, K. A. (2020). Classification of Poverty Levels Using Machine Learning. *Journal of Xi'an University of Architecture & Technology*, 12(4), 5723-5728.
- Sihombing, P. R., & Arsani, A. M. (2021). Comparison of Machine Learning Methods in Classifying Poverty in Indonesia in 2018. *Jurnal Teknik Informatika (Jutif)*, 2(1), 51-56. <https://doi.org/10.20884/1.jutif.2021.2.1.52>
- Rozenberg, J., & Hallegatte, S. (2016). Model and Methods for Estimating the Number of People Living in Extreme Poverty Because of the Direct Impacts of Natural Disasters. *Model and Methods for Estimating the Number of People Living in Extreme Poverty Because of the Direct Impacts of Natural Disasters*, November. <https://doi.org/10.1596/1813-9450-7887>
- Mustapha Ibrahim, D., & Ren, D. (2021). Nigeria Poverty Analysis, Prediction Using Machine Learning Methods and Deep Learning. *Asian Journal of Social Sciences, Arts and Humanities*, 9(1), 2021. [www.multidisciplinaryjournals.com](http://www.multidisciplinaryjournals.com)
- Alsharkawi, A., Al-Fetyani, M., Dawas, M., Saadeh, H., & Alyaman, M. (2021). Poverty classification using machine learning: The case of Jordan. *Sustainability (Switzerland)*, 13(3), 1-16. <https://doi.org/10.3390/su13031412>
- Ayush, K., Uz Kent, B., Tanmay, K., Burke, M., Lobell, D., & Ermon, S. (2020). Efficient Poverty Mapping using Deep Reinforcement Learning. <http://arxiv.org/abs/2006.04224>
- Huang, L. Y., Hsiang, S., & Gonzalez-Navarro, M. (2021). Using Satellite Imagery and Deep Learning to Evaluate the Impact of Anti-Poverty Programs. *SSRN Electronic Journal*, 1-37. <https://doi.org/10.2139/ssrn.3897541>
- Jiang, Y., Zhang, L., Li, Y., Lin, J., Li, J., Zhou, G., Liu, S., Cao, J., & Xiao, Z. (2021). Evaluation of county-level poverty alleviation progress by deep learning and satellite observations. *Big Earth Data*, 5(4), 576-592
- The SMERU Research Institute. (2016). Penetapan Kriteria dan Variabel Pendataan Penduduk Miskin yang Komprehensif dalam Rangka Perlindungan Penduduk Miskin di Kabupaten/Kota.