

## Penerapan Data Mining Untuk Klasterisasi Tingkat Kemiskinan Berdasarkan Data Terpadu Kesejahteraan Sosial (DTKS)

Muhammad Ihza Zuhendra<sup>1</sup>, Rahmad Hidayat<sup>2\*</sup>, Hendrawaty<sup>3</sup>

<sup>1,2,3</sup>Teknologi Informasi dan Komputer, Teknik Informatika, Politeknik Negeri Lhokseumawe, Lhokseumawe, Indonesia

Email: <sup>1</sup>muhammadiyah2001@gmail.com, <sup>2</sup>rahmad\_hidayat@pnl.ac.id, hendrawaty@pnl.ac.id

(\* : corresponding author)

### Abstrak

Tingkat kemiskinan menjadi nilai yang mempengaruhi kesejahteraan suatu negara. Kemiskinan dapat disebabkan oleh beberapa faktor seperti terbatasnya lowongan pekerjaan sehingga minimnya penghasilan yang didapat namun tidak cukup untuk menopang biaya hidup, banyaknya tanggungan keluarga dan sebagainya. Dalam hal ini, pemerintah ikut berperan dengan memberikan bantuan seperti bantuan sosial. Tahapan dalam pemberian bantuan ini, salah satunya warga diharuskan terdaftar sebagai peserta Data Terpadu Kesejahteraan Sosial (DTKS). Menjadi salah satu peserta DTKS harus memenuhi persyaratan terkategori sebagai fakir miskin yang secara umum ditentukan melalui keputusan menteri sosial nomor: 146/HUK/2013 tentang kriteria fakir miskin teregister. Pertimbangan ini dapat dijadikan panduan dalam menentukan status kelompok masyarakat dalam suatu wilayah. Namun, dalam skala yang luas, mengelompokkan masyarakat berdasarkan tingkat kemiskinan dapat menjadi tugas yang kompleks dan memakan waktu. K-Means merupakan salah satu dari beberapa metode data clustering non hirarki dengan sistem kerja mempartisi data yang ada ke dalam bentuk satu atau lebih klaster/kelompok. Klasterisasi ini dapat diterapkan dalam mengelompokkan sekumpulan data besar untuk meningkatkan keakuratan informasi yang didapatkan, contohnya untuk mengetahui tingkat kemiskinan suatu daerah. Tujuan penelitian ini untuk membangun aplikasi yang memudahkan untuk melakukan pengelompokkan masyarakat dalam melihat perkembangan tingkat kemiskinan pada suatu wilayah berdasarkan kriteria yang telah ditetapkan sehingga membantu pemerintah dan pemangku kepentingan lainnya untuk lebih memahami distribusi kemiskinan, mengidentifikasi kelompok yang berisiko tinggi dan merancang program bantuan sosial atau kebijakan yang lebih terarah dan efektif. Hasil dari penelitian ini menampilkan visualisasi dari presentase tiap-tiap kelompok pada suatu data. Visualisasi data yang disajikan juga dapat ditampilkan berdasarkan kategori seperti jumlah klaster, wilayah, tahun dan lainnya.

Kata Kunci: Klasterisasi, Tingkat Kemiskinan, K-Means, DTKS

### Abstract

*The level of poverty serves as a significant indicator influencing a nation's well-being. Poverty can arise from various factors, such as limited job opportunities resulting in insufficient income to cover living expenses, substantial family responsibilities, and more. In this context, the government plays a role by providing assistance, such as social aid programs. One step in providing this assistance involves individuals being registered as participants in the Unified Social Welfare Data (Data Terpadu Kesejahteraan Sosial or DTKS). Becoming a DTKS participant requires meeting the criteria categorizing someone as extremely poor, which is generally determined by the Minister of Social Affairs' Decision No. 146/HUK/2013 on the criteria for registered individuals in extreme poverty. This consideration can serve as a guideline in determining the socioeconomic status of community groups within a region. However, on a larger scale, classifying communities based on poverty levels can be a complex and time-consuming task. K-Means clustering is one of several non-hierarchical data clustering methods that work by partitioning existing data into one or more clusters or groups. This clustering can be applied to categorize a large dataset to enhance the accuracy of the information obtained, such as assessing the poverty level in a specific area. The objective of this research is to develop an application that facilitates the categorization of communities in analyzing the progression of poverty rates in a region based on predefined criteria. This aids the government and other stakeholders in understanding poverty distribution better, identifying high-risk groups, and designing targeted and effective social aid programs or policies. The outcomes of this research showcase visualizations depicting the percentage composition of each group within a dataset. The presented data visualizations can also be customized based on categories such as the number of clusters, regions, years, and more.*

*Keywords: Clusterization, Poverty Rate, K-Means, DTKS*

## 1. PENDAHULUAN

Kemiskinan merupakan salah satu masalah sosial yang masih menjadi tantangan besar bagi negara kita karena dapat menimbulkan ketimpangan sosial yang hal ini akan menyebabkan orang yang kaya semakin kaya dan yang miskin semakin miskin [1]. Kemiskinan dapat disebabkan oleh beberapa faktor seperti terbatasnya lowongan pekerjaan sehingga minimnya penghasilan yang didapat tidak cukup menopang biaya hidup, banyaknya tanggungan keluarga dan sebagainya [2]. Perekonomian yang cenderung menurun mengakibatkan jumlah penduduk miskin yang meningkat. Jumlah penduduk miskin di Indonesia tren nya cenderung menurun akan tetapi Maret 2020 mengalami peningkatan [3].

Kemiskinan adalah terciptanya rasa ketidakberdayaan yang mendorong adanya kerentanan menghadapi situasi dalam dirinya sendiri sehingga memunculkan ketergantungan dan keterasingan baik secara geografis maupun sosiologis. Kondisi di mana seseorang atau kelompok tidak memiliki atau tidak dapat mengakses sumber daya ekonomi, sosial, dan politik yang memadai (World Bank). Kemiskinan merupakan masalah yang tidak pernah lepas dari perhatian pemerintah di berbagai belahan negara. Kemiskinan juga menjadi permasalahan besar dalam bidang ekonomi yang menjadi titik acuan keberhasilan suatu negara dari waktu ke waktu, terutama pada negara-negara yang sedang berkembang. Indonesia sebagai salah satu negara berkembang menyadari pentingnya memperhatikan masalah ini upaya menekan laju kemiskinan [4].

Tingkat kemiskinan menjadi suatu nilai yang mempengaruhi kesejahteraan suatu negara. Tingkat kemiskinan adalah presentase jumlah penduduk yang berada dibawah garis kemiskinan, yaitu minimum untuk memperoleh standar hidup yang mencukupi di suatu negara [5]. Dalam hal ini, pemerintah ikut berperan dengan memberikan bantuan seperti bantuan sosial (bansos), sembako dan sebagainya. Bantuan ini memiliki beberapa tahapan dalam pemberiannya salah satunya warga yang mendapatkan bantuan sosial harus terdaftar sebagai peserta Data Terpadu Kesejahteraan Sosial (DTKS). DTKS meliputi Pemerlu Pelayanan Kesejahteraan Sosial (PPKS), Penerima Bantuan dan Pemberdaan Sosial serta Potensi dan Sumber Kesejahteraan Sosial (PSKS) [6]. Masyarakat yang ingin atau belum terdaftar dalam peserta DTKS, dapat melakukannya melalui pendaftaran mandiri DTKS Kemensos. Untuk menjadi salah satu peserta DTKS harus memenuhi persyaratan sebagai fakir miskin. Penetapan kriteria fakir miskin ini secara umum ditentukan melalui keputusan menteri sosial nomor: 146/HUK/2013 tentang kriteria fakir miskin teregister [7].

Banyaknya kriteria dan jumlah masyarakat calon penerima bantuan menjadikan nilai-nilai penting yang patut dipertimbangkan. Pertimbangan ini dapat dijadikan panduan dalam menentukan status suatu kelompok masyarakat atau suatu wilayah. Penentuan kelompok – kelompok ini ditujukan dalam memprioritaskan pemberian bantuan sosial sehingga dapat menekan peningkatan pertumbuhan angka kemiskinan. Dalam hal ini maka data yang digunakan tergolong pada data yang terbilang cukup banyak atau besar, sehingga penelitian ini termasuk ke dalam proses data mining yang mencari informasi berdasarkan banyak data. Proses pencarian dan penggalian informasi dari tumpukan-tumpukkan data dengan jumlah yang besar merupakan proses utama dari data mining, tujuan utama dari pengolahan data tersebut adalah menghasilkan informasi yang baru [8]. Data mining memiliki segudang manfaat untuk mengelola data mentah menjadi knowledge atau kumpulan informasi yang dapat dijadikan sebagai penunjang keputusan secara efektif [9].

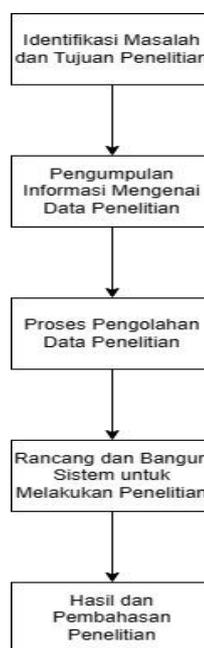
Analisis klastering merupakan suatu teknik pada data mining yang bertujuan untuk mengidentifikasi suatu kelompok dari objek yang memiliki karakteristik yang sama. K-means merupakan algoritma klastering data yang dibagi ke dalam satu atau lebih klaster yang memiliki karakteristik sama dihasilkan hasil pengelompokan berdasarkan kriteria tertentu [10].

Dari permasalahan diatas maka diangkatlah sebuah judul penelitian “Penerapan Data Mining Untuk Klasterisasi Tingkat Kemiskinan Berdasarkan Data Terpadu Kesejahteraan Sosial (DTKS) Menggunakan Metode K-Means”, upaya membantu mengatasi persoalan dalam menentukan status pada suatu kelompok masyarakat. Adapun tujuan penelitian ini yaitu untuk

membuat aplikasi yang memudahkan untuk melakukan pengelompokan masyarakat dalam melihat perkembangan tingkat kemiskinan pada suatu wilayah berdasarkan kriteria yang telah ditetapkan. Selain itu aplikasi ini juga dapat berguna bagi Badan Pusat Statistik (BPS) kota Lhokseumawe untuk mendapatkan informasi berdasarkan hasil klasterisasi suatu data dalam jumlah yang besar. Diharapkan hasil dari penelitian ini dapat memberikan aplikasi yang memudahkan dalam melihat perkembangan status kemiskinan sehingga membantu pemerintah dan pemangku kepentingan lainnya untuk lebih memahami distribusi kemiskinan, mengidentifikasi kelompok yang berisiko tinggi dan merancang program bantuan sosial atau kebijakan yang lebih terarah dan efektif.

## 2. METODE PENELITIAN

Secara sederhana alur tahapan pembuatan sistem untuk penelitian yang dilakukan dalam pembuatan sistem dapat diimplementasikan seperti yang terlihat pada Gambar 1 berikut:



Gambar 1. Alur Tahapan Penelitian

Tahapan alur dalam penelitian ini meliputi:

### a. Latar Belakang Masalah,

Tahapan ini merupakan tahapan dalam memulai suatu penelitian dan tujuan dalam melakukan penelitian. Pada penelitian ini, latar belakang masalah yang didapat yaitu bagaimana melakukan klasterisasi dengan mempertimbangkan kriteria DTKS yang dijadikan panduan dalam menentukan status kelompok masyarakat dalam suatu wilayah. Namun, dalam skala yang luas, mengelompokkan masyarakat berdasarkan tingkat kemiskinan dapat menjadi tugas yang kompleks dan memakan waktu. Kemudian dengan permasalahan yang ada, didapat tujuan penelitian untuk membangun aplikasi yang memudahkan untuk melakukan pengelompokan masyarakat dalam melihat perkembangan tingkat kemiskinan pada suatu wilayah berdasarkan kriteria yang telah ditetapkan sehingga membantu pemerintah dan pemangku kepentingan lainnya untuk lebih memahami distribusi kemiskinan, mengidentifikasi kelompok yang berisiko tinggi dan merancang program bantuan sosial atau kebijakan yang lebih terarah dan efektif.

### b. Pengumpulan Data

Dalam Tahapan ini yang dilakukan dengan mewawancarai badan pemerintah yang bersangkutan seperti Dinas Sosial (Dinsos) dan Badan Pusat Statistik (BPS). Pembahasan dalam

wawancara mengenai kriteria-kriteria yang digunakan dalam menentukan kemiskinan pada suatu wilayah berdasarkan sesuai hak dan wewenang pada badan setempat. Kemudian mencari informasi mengenai saran dalam pengelompokan yang dilakukan, sehingga dapat digunakan untuk penelitian ini maupun penelitian berikutnya.

c. Pengolahan Data

Pada proses pengolahan data disini yaitu dengan melakukan pemberian nilai bobot khusus pada kriteria-kriteria yang bertipe karakter. Hal ini bertujuan agar metode K-Means dapat diimplementasikan secara maksimal.

d. Rancang dan Bangun Sistem

Perancangan dan bangun sistem dilakukan untuk mengidentifikasi fitur apa saja untuk memenuhi kebutuhan sistem dalam mengimplementasikan penelitian ini ke dalam sistem sehingga mendapatkan hasil sesuai dengan tujuan dari penelitian yang dilakukan.

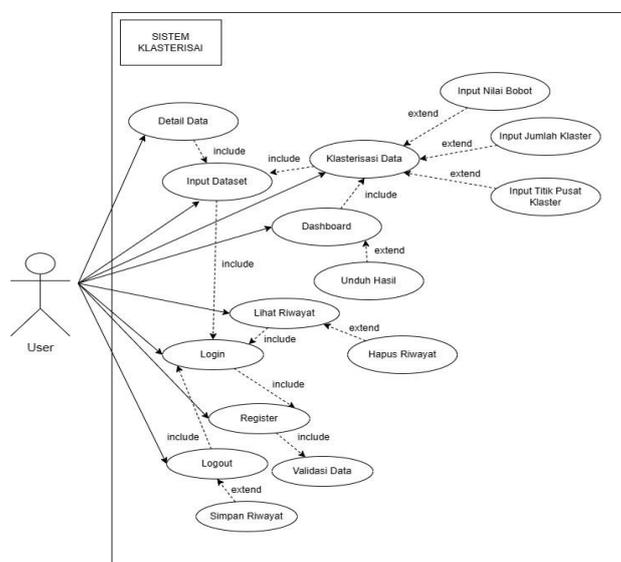
e. Hasil dan Pembahasan

Tahapan ini berguna untuk mengetahui proses dari pengerjaan dan suatu penelitian dan dapat diketahui apakah sistem dapat melakukan visualisasi dari hasil pengklasteran data menggunakan metode K-Means nantinya.

2.1 Metodologi Penelitian

a. Use Case Diagram

Berdasarkan perancangan *use case diagram*, pada sistem terdapat tahapan-tahapan yang dapat dilakukan *user* pada sistem untuk melakukan klasterisasi pada data. *Use case diagram* dapat dilihat seperti Gambar 2 berikut:



Gambar 2. Use Case Diagram Klasterisasi Tingkat Kemiskinan

Pada Gambar 2 *Use Case Diagram* terdapat satu *actor* yaitu *user* yang dapat mengakses sistem dengan uraian tahapan-tahapannya sebagai berikut:

- 1) Register. Pada *use case register*, dapat melakukan registrasi akun pada sistem untuk dapat login.
- 2) Login. Pada *use case user login*, *user* dapat *login* ke dalam sistem setelah melalui proses berupa registrasi

- 3) Input DataSet. Pada *use case* input dataset, dapat memasukkan sebuah inputan berupa input dataset ke dalam sistem setelah *user* melakukan *login*. Dataset ini akan digunakan untuk melakukan proses klasterisasi pada tahapan berikutnya.
- 4) Detail Data. Pada *use case* detail data, aktifitas yang dapat dilakukan oleh *user* yaitu dapat melihat detail data dari inputan dataset yang telah diinput melalui halaman Input Dataset.
- 5) Klasterisasi Data. Pada halaman klasterisasi, aktifitas yang dilakukan oleh *user* yaitu melakukan klasterisasi. Klasterisasi dapat dilakukan setelah *user* melakukan inputan berupa sebuah dataset ke dalam sistem pada halaman Input Dataset. Kemudian untuk melanjutkan proses klasterisasi, *user* dapat menentukan nilai bobot dengan cara menginputkan nilai bobot. Kemudian *user* juga dapat menentukan jumlah klaster. Selanjutnya *user* dapat menetapkan titik pusat untuk tiap klaster berdasarkan jumlah klaster yang telah ditentukan.
- 6) Dashboard. Pada halaman dashboard, *user* dapat melihat hasil klasterisasi dalam bentuk visualisasi data setelah melakukan proses klasterisasi.
- 7) Lihat Riwayat. Pada halaman riwayat, *user* dapat melihat riwayat akses ke dalam sistem.
- 8) *Logout*. *User* dapat keluar dari sistem dengan menekan tombol *logout* pada sistem dan kemudian akan menyimpan tanggal *user* akses sistem ke riwayat.

#### b. Metode K-Means

Klaster mengacu pada kumpulan titik data yang dikumpulkan bersama karena kesamaan tertentu. Jika didefinisikan  $K = 2$ , maka akan ada 2 klaster, jika  $K = 3$  maka terdapat 3 klaster dan begitu seterusnya. Dengan demikian K-means klastering dapat didefinisikan sebagai algoritma iteratif yang membagi kumpulan data (dataset) yang tidak berlabel menjadi 'k' klaster yang berbeda sedemikian rupa sehingga setiap kumpulan data hanya dimiliki oleh satu kelompok yang memiliki properti serupa.

Secara umum metode K-Means klaster analisis menggunakan algoritma sebagai berikut:

- 1) Menyiapkan dataset. Tahap awal dengan menyiapkan data yang akan dikelompokkan baik setelah melalui tahapan data processing.
- 2) Tentukan k sebagai jumlah cluster yang di bentuk. Untuk menentukan berapa banyak *cluster* dilakukan dengan pertimbangan teoritis dan konseptual yang kemungkinan akan diusulkan.
- 3) Tetapkan k Centroid (titik pusat cluster) awal secara acak. Penentuan *centroid* awal dilakukan secara acak berdasarkan objek-objek yang tersedia sebanyak k *cluster*, setelah itu melakukan perhitungan centroid cluster ke-i berikutnya, menggunakan rumus:

$$v = \frac{\sum_{i=1}^n X_i}{n} \quad (1)$$

dengan:

$v$  : centroid pada *cluster*

$x_i$ : objek ke-i n

$n$ : banyaknya objek/jumlah objek yang menjadi anggota *cluster*

- 4) Menghitung jarak setiap objek ke masing-masing centroid yang telah ditetapkan untuk masing-masing cluster. Menghitung jarak antara objek dengan centroid menggunakan Euclidian Distance

$$d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (2)$$

dengan :

- $x_i$ : objek  $x$  ke- $i$
- $y_i$ : objek  $y$  ke- $i$
- $n$ : banyaknya objek

- 1) Alokasikan masing-masing objek ke dalam centroid yang paling terdekat.
- 2) Lakukan iterasi, kemudian menentukan posisi centroid baru dengan menggunakan persamaan (1). Ulangi langkah 3 jika posisi centroid baru tidak sama.

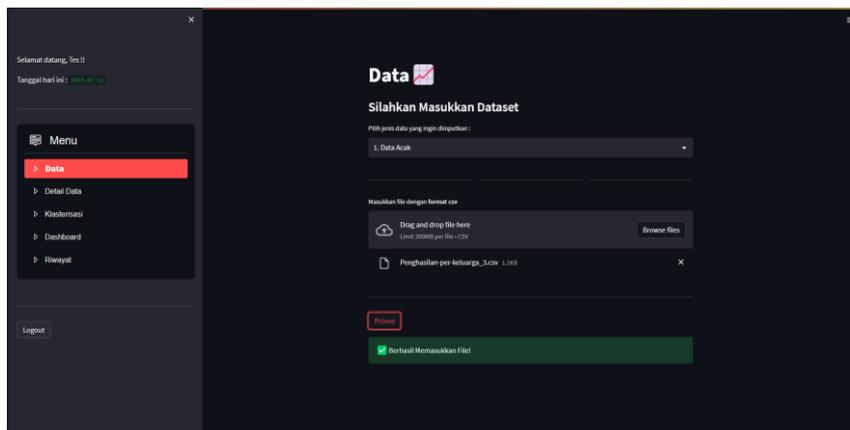
### 3. HASIL DAN PEMBAHASAN

Hasil perancangan sistem yang telah dibuat untuk menampilkan hasil visualisasi dari pengelompokan data menggunakan metode K-Means dibahas pada sub bab sebagai berikut.

#### 3.1 Tampilan *User Interface*

##### a. Halaman Data

Pada halaman data, *user* dapat melakukan proses inputan berupa data yang akan dilakukan klusterisasi menggunakan metode K-Means. *User interface* halaman data dapat dilihat seperti Gambar 3 berikut:

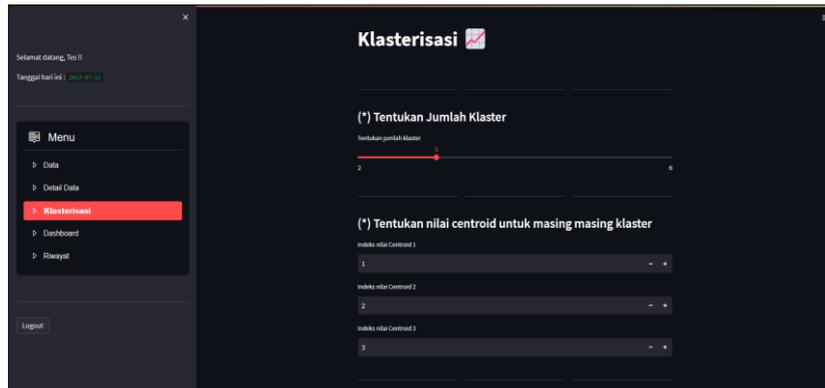


Gambar 3. Halaman Data

Halaman data merupakan halaman muka, ataupun halaman awal pada sistem. Pada halaman data, *user* dapat melakukan inputan berupa dataset, nantinya data inilah yang akan digunakan untuk melakukan proses klusterisasi. Pada gambar kita dapat melihat sebuah opsi, dimana kita dapat menginputkan data berdasarkan dua jenis data, yaitu data acak atau data kemiskinan. Ketika ingin menginputkan data kemiskinan maka data ini merujuk dengan data data yang memiliki atribut yang sesuai berdasarkan kategori kemiskinan. Sedangkan data acak, data ini tidak merujuk dengan suatu atribut khusus. Kemudian, *user* juga dapat menginputkan data dengan menekan tombol 'browse files' dari *device* computer *user*. Selanjutnya untuk memasukkan data, *user* hanya perlu menekan tombol proses diikuti dengan pesan berhasil memasukkan data yang akan diberitahukan oleh sistem.

##### b. Halaman Klusterisasi

Pada halaman klusterisasi, *user* dapat melakukan proses pengelompokan. Pengelompokan yang dilakukan menggunakan metode K-Means. *User interface* halaman klusterisasi dapat dilihat seperti pada Gambar 4 berikut:



Gambar 4. Halaman Klasterisasi

Gambar 4 menampilkan halaman klasterisasi pada sistem. Pada halaman klasterisasi, beberapa fitur yang dapat digunakan oleh user yaitu:

1) *User* dapat menentukan jumlah klaster

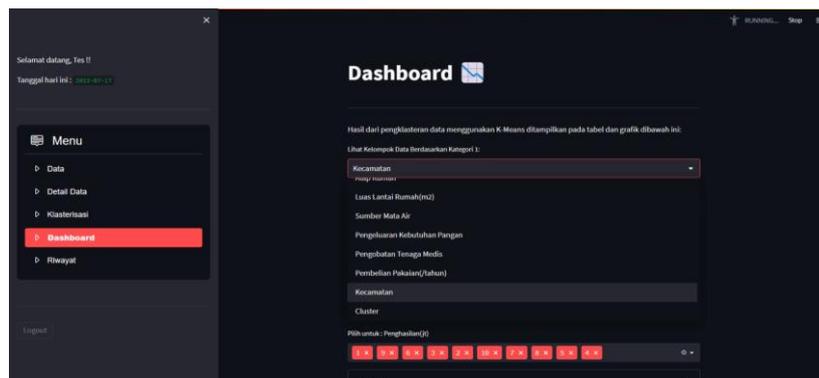
Penentuan jumlah klaster akan mempengaruhi bentuk output dari visualisasi data yang dilihat nantinya. Misalkan memilih untuk 3 klaster, maka data kemudian akan diklasterisasikan menjadi 3 kelompok, apabila 4 klaster maka data akan diklasterisasikan menjadi 4 kelompok dan seterusnya.

2) *User* dapat menentukan nilai centroid untuk tiap tiap klaster.

Penentuan nilai centroid ini dipacu berdasarkan jumlah klaster yang ditentukan sebelumnya. Dalam hal ini apabila klaster yang dipilih sebanyak 2 klaster, maka diperlukan untuk menentukan nilai centroid sebanyak 2 titik. Apabila dipilih 3 klaster, maka diperlukan untuk menentukan nilai centroid sebanyak 3 titik dan seterusnya.

### c. Halaman Dashboard

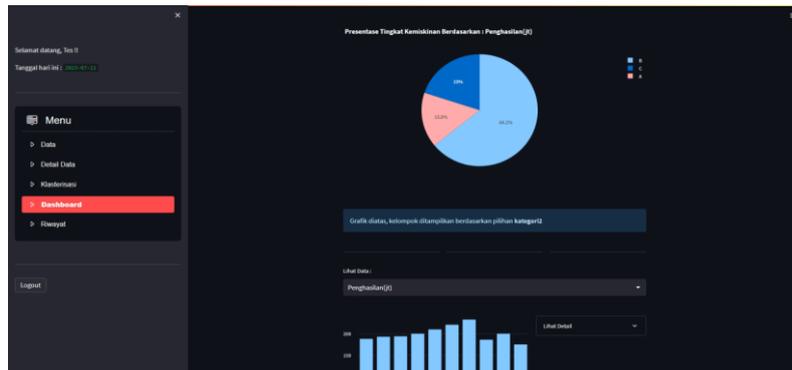
Pada halaman dashboard, *user* dapat melihat hasil dari klasterisasi data berdasarkan pengelompokan yang dilakukan menggunakan metode K-Means. *User interface* halaman dashboard dapat dilihat pada Gambar 5 berikut:



Gambar 5. Halaman Dashboard Pemilihan Kategori

Halaman dashboard menjadi halaman dimana sistem akan menampilkan visualisasi dari hasil klasterisasi data inputan user. Hasil visualisasi ini tersedia dalam bentuk grafik dan juga tabel. Pada gambar halaman dashboard diatas, user dapat memilah hasil visualisasi *output* berdasarkan 2 kategori. Kategori ini didasarkan pada atribut-atribut yang ada pada data yang

dimasukkan oleh user. Misalkan pada contoh di Gambar 5, *output* dapat dipilah berdasarkan Kecamatan dan Penghasilan.



Gambar 6. Halaman Dashboard Tampilan Visualisasi Data

Kemudian setelah ditentukan kategori yang divisualisasikan berdasarkan data yang ada, maka sistem akan menampilkannya ke dalam bentuk *Pie Chart* seperti gambar di atas. Dapat dilihat jumlah presentasi dibagi menjadi 3 warna. Setiap warna mewakili jumlah presentasi untuk tiap tiap klaster. Kemudian sistem juga dapat menampilkan detail visualisasi data berdasarkan kategori yang dipilih, misalkan pada Gambar 6 kategori yang dipilih berupa ‘Penghasilan(jt)’. Selanjutnya sistem akan merepresentasikannya ke dalam bentuk *Bar Chart*. Pada tombol ‘Lihat Detail’ *user* dapat melihat nilai tertinggi dan terendah serta rata-rata berdasarkan atribut yang dipilih sebelumnya.

Perhitungan manual dijadikan suatu alternatif untuk melihat apakah implementasi algoritma K-Means pada sistem sudah sesuai. Untuk melihatnya, disini diuji dengan membandingkan hasil klasterisasi data menggunakan perhitungan manual dengan hasil *output* klasterisasi data menggunakan sistem.

Tabel 1. Data Sampel Pengujian

X	Y
10,9	12,6
2,3	8,4
8,4	12,6
12,1	16,2
7,3	8,9
23,4	11,3
19,7	18,5
17,1	17,2
3,2	3,4
1,3	22,8
2,4	6,9
2,4	7,1
3,1	8,3
2,9	6,9
11,2	4,4
8,3	8,7
10,7	12,2
2,1	8,2
8,2	12,4
11,9	16
7,1	8,7
23,2	11,1
19,5	18,3
16,9	17
3,0	3,2

Data yang disajikan pada Tabel 1 merupakan data sampel yang digunakan untuk melakukan pengujian pada sistem. Pengujian dilakukan dengan membandingkan hasil klusterisasi berdasarkan data sampel Tabel 1 menggunakan perhitungan manual dan hasil klusterisasi berdasarkan data sampel tersebut menggunakan sistem. Data sampel diatas memiliki dua atribut yaitu atribut x dan atribut y, dengan panjang data 25 sampel data. Kemudian dilakukan perhitungan manual menggunakan excel.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
		centroid1		centroid2		centroid3				dc1		dc2		dc3		class						
0	10,9	12,6	23,4	11,3	3,2	3,4	10,9	12,6	0	12,57	12,00				1	10,9	12,6					
1	2,3	8,4	5	8			2,3	8,4	9,57	21,30	5,08				3					2,3	8,4	
2	8,4	12,6					8,4	12,6	2,50	15,06	10,57				1	8,4	12,6					
3	12,1	16,2					12,1	16,2	3,79	12,32	15,59				1	12,1	16,2					
4	7,9	8,9					7,9	8,9	5,16	16,28	4,86				1	7,9	8,9					
5	23,4	11,3					23,4	11,3	12,57	0,00	21,69				2			23,4	11,3			
6	19,7	18,5					19,7	18,5	10,59	8,10	22,37				2			19,7	18,5			
7	17,1	17,2					17,1	17,2	7,72	8,63	19,59				1	17,1	17,2					
8	3,2	3,4					3,2	3,4	12,00	21,69	0,00				3					3,2	3,4	
9	1,3	22,8					1,3	22,8	14,01	24,91	19,49				1	1,3	22,8					
10	2,4	6,9					2,4	6,9	10,23	21,46	3,59				3					2,4	6,9	
	2,4	7,1					2,4	7,1	10,12	21,42	3,79				3					2,4	7,1	
	3,1	8,3					3,1	8,3	8,91	20,52	4,90				3					3,1	8,3	
	2,9	6,9					2,9	6,9	9,82	20,97	3,51				3					2,9	6,9	
	11,2	4,4					11,2	4,4	8,21	14,02	8,06				3					11,2	4,4	
	8,3	8,7					8,3	8,7	4,69	15,33	7,36				1	8,3	8,7					
	10,7	12,2					10,7	12,2	0,45	12,73	11,56				1	10,7	12,2					
	2,1	8,2					2,1	8,2	9,84	21,52	4,92				3					2,1	8,2	
	8,2	12,4					8,2	12,4	2,71	15,24	10,30				1	8,2	12,4					
	11,9	16					11,9	16	3,54	12,43	15,31				1	11,9	16					
	7,1	8,7					7,1	8,7	5,45	16,51	6,58				1	7,1	8,7					
	23,2	11,1					23,2	11,1	12,39	0,28	21,43				2			23,2	11,1			
	19,5	18,3					19,5	18,3	10,32	8,01	22,08				2			19,5	18,3			
	16,9	17					16,9	17	7,44	8,65	19,30				1	16,9	17					
	3	3,2					3	3,2	12,28	21,95	0,28				3					3	3,2	
																10,02	13,78	21,45	14,80	3,62	6,31	

Gambar 7. Hasil Perhitungan Manual Menggunakan Excel

Pada Gambar 7 merupakan gambar hasil klusterisasi melalui excel berdasarkan algoritma klusterisasi K-Means. Pada hasil perhitungan manual tersebut dapat dilihat terdapat 25 data sampel dan 3 titik *centroid*, yang berarti, pada data diatas dilakukan klusterisasi sebanyak 3 kluster. Sehingga jumlah dari *centroid* harus sebanyak 3 titik *centroid*. Berdasarkan gambar, titik *centroid* untuk kluster 1 diambil berdasar data dengan indeks ke-0 atau elemen ke-1, titik *centroid* untuk kluster 2 diambil berdasarkan data dengan indeks ke-5 atau elemen ke-6 dan titik *centroid* untuk kluster 3 diambil berdasarkan data dengan indeks ke-8 atau elemen ke-9 pada data sampel. Kemudian berdasarkan pusat kluster tersebut, maka dihitung untuk jarak setiap data terhadap *centroid* menggunakan rumus (1). Kemudian data dapat ditentukan termasuk ke dalam kluster berdasarkan nilai terendah yang didapat setelah dilakukan perhitungan untuk tiap data terhadap pusat kluster seperti pada gambar. Langkah selanjutnya adalah dilakukan pencarian nilai titik *centroid* baru menggunakan rumus (2). Apabila terjadi pergeseran titik *centroid* maka perlu dilakukan iterasi baru, namun menggunakan titik *centroid* yang baru hingga tidak lagi mengalami perubahan.

Selanjutnya, dengan sampel data yang sama, dilakukan klusterisasi namun menggunakan bantuan sistem. Pada sistem ditentukan jumlah kluster sebanyak 3 kluster dan 3 titik *centroid* berdasarkan indeks pada data ke-0, ke-5 dan ke-8. Didapati hasilnya sebagai berikut:

class	centroid
1	1
3	3
1	1
1	1
3	3
2	2
2	2
3	3
1	1
3	3
3	3
3	3
3	3
3	3
1	1
3	3
1	1
1	1
3	3
2	2
2	2
2	2
3	3

Gambar 8. Hasil Perbandingan Klusterisasi Perhitungan Manual(kiri) dan Sistem(kanan)

Pada Gambar 8 merupakan gambar hasil dari klasterisasi menggunakan perhitungan manual dan klasterisasi menggunakan bantuan sistem. Pada gambar dapat dilihat bahwa hasil dari perhitungan manual dan sistem memiliki *output* yang sama, dengan jumlah sampel data yang sama.

#### 4. KESIMPULAN

Setelah adanya dilakukan perancangan, pembuatan dan pengujian pada aplikasi sistem klasterisasi tingkat kemiskinan menggunakan algoritma K-Means, maka dapat diambil kesimpulan, yaitu hasil akhir pada sistem ketika dilakukan uji coba, sistem pada penelitian ini dapat menampilkan visualisasi dari presentase tiap-tiap kelompok pada suatu data. Visualisasi data yang disajikan juga dapat ditampilkan berdasarkan kategori seperti jumlah klaster, wilayah, tahun dan lainnya. Berdasarkan perbandingan proses klasterisasi menggunakan metode K-Means antara perhitungan manual dan perhitungan menggunakan sistem, dapat diambil kesimpulan berdasarkan uji coba kinerja sistem memiliki akurasi yang tinggi dalam melakukan proses klasterisasi.

#### UCAPAN TERIMA KASIH

Setelah melakukan penelitian ini, peneliti mengucapkan terima kasih sebesar-besarnya kepada:

- 1) Kepada lembaga pendidikan Politeknik Negeri Lhokseumawe yang telah berkontribusi dalam melakukan pengarahan dan pengalaman Pendidikan.
- 2) Kepada *staff* dan dosen-dosen Politeknik Negeri Lhokseumawe yang telah membina dan memberikan arahan.
- 3) Kepada bapak Dr. Rahmad Hidayat, S.Kom., M.CS dan ibu Hendrawaty. ST., MT. yang sudah meluangkan waktu untuk dapat membimbing dalam melakukan penelitian.
- 4) Kepada anggota keluarga yang selalu memberikan dukungan dan semangat dalam setiap kondisi.
- 5) Kepada kerabat dan teman-teman seperjuangan.

#### DAFTAR PUSTAKA

- [1] Ezra, V. M., and Nandika A. L., "Strategi Pemanfaatan Data Terpadu Kesejahteraan Sosial (DTKS) Dalam Penyaluran Bantuan Sosial RS-RTLH Oleh Dinas Sosial Provinsi Sulawesi Utara," *Jurnal Konstituen.*, vol. 4, no. 1, pp. 25-39, 2022.
- [2] Kausar, A. G., "Analisis Pengaruh Tingkat Pengangguran Terbuka, Kesempatan Kerja Dan Tingkat Pendidikan Terhadap Tingkat Kemiskinan Di Kabupaten Aceh Barat," *Skripsi, Ekonomi, Universitas Teuku Umar, Meulaboh, Indonesia*, 2022.
- [3] Nugroho, A. S., et al., "Penerapan Data Mining Pada Jumlah Penduduk Miskin Di Indonesia," *Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika.*, vol. 1, no. 3, pp. 199-207, 2020.
- [4] Hilmi, H., et al., "Pengaruh Jumlah Penduduk dan Pengangguran Terhadap Tingkat Kemiskinan Di Kabupaten Tolitoli," *Jurnal Ilmiah Ekonom Pembangunan.*, vol. 1, no. 1, pp. 20-27, 2022.
- [5] Ari, S., and Eko, S., "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *Jurnal Tekno Kompak.*, vol. 15, no. 2, pp. 25-36, 2020.
- [6] Muhammad, I., "Edukasi Pengelolaan Keuangan Keluarga Penerima Manfaat Program Penerima Harapan Di Kelurahan Nali Kecamatan Baolan Kabupaten Tolitoli," *Jurnal Pengabdian Kepada Masyarakat.*, vol. 2, no.3, pp. 743-750, 2022.
- [7] (2022), website.desa.id. [online]. Available: <https://3318032013.website.desa.id/>
- [8] Rizki, M., and Zulfikar, S., "Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen," *CESS (Journal of Computer Engineering System and Science)*, vol. 4, no. 2, pp. 272-279, 2019.

- [9] Yunita, R. S., et al, “Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer,” *Journal of Computer Engineering System and Science.*, vol. 5, pp. 192-198, 2020.
- [10] Tuti, H., Odi, N., Eko, W., “Analisis dan Penerapan Algoritma K-Means Dalam Strategi Promosi Kampus Akademi Maritim Suaka Bahari,” *Jurnal Sains Teknologi Transportasi Maritim.*, vol. 3, pp. 1-7, 2021.