

**Penerapan Algoritma *Fuzzy C-Means* dan Metode *Elbow* untuk  
Mengelompokkan Provinsi di Indonesia Berdasarkan Indeks Demokrasi Indonesia**

***Application of The Fuzzy C-Means Algorithm and Elbow Method to  
Group Provinces in Indonesia Based on The Indonesian Democracy Index***

**Nina Nurdiana<sup>1</sup>, Agung Nilogiri<sup>2\*</sup>, Miftahur Rahman<sup>3</sup>**

<sup>1</sup>Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [ninanurdiana1299@gmail.com](mailto:ninanurdiana1299@gmail.com)

<sup>2</sup>Dosen Fakultas Teknik, Universitas Muhammadiyah Jember\* Koresponden Author  
email: [agungnilogiri@unmuhjember.ac.id](mailto:agungnilogiri@unmuhjember.ac.id)

<sup>3</sup>Dosen Fakultas Teknik, Universitas Muhammadiyah Jember  
email: [miftahurrahman@unmuhjember.ac.id](mailto:miftahurrahman@unmuhjember.ac.id)

**Abstrak**

Negara Indonesia ialah sebuah negara berdasarkan demokrasi. Adapun pernyataan Abraham Lincoln, mengatakan bahwa pelaku dari pemerintahan adalah rakyat. Pihak terkait melakukan pengukuran kemajuan demokrasi dengan model pengukuran Indeks Demokrasi Indonesia (IDI). Tujuan dari IDI adalah membaca pencapaian dan menyusun program-program perencanaan membangun politik. *Clustering* adalah suatu proses pembagian data dari sebuah himpunan dalam beberapa kelompok dimana data satu dengan data yang lain jika mempunyai kemiripan lebih dominan akan membentuk sebuah kelompok, dan data yang lain akan dipisahkan dengan membentuk kelompok yang berbeda menjadi kelompok yang berbeda. Penelitian ini dilakukan untuk mengetahui informasi dan pengetahuan baru mengenai penyebab terjadinya penurunan skor berdasarkan data Indeks Demokrasi Indonesia tahun 2019 dan untuk menemukan kluster terbaik dalam *clustering* Provinsi di Indonesia berlandaskan dataset dari Indeks Demokrasi Indonesia 2019. *Clustering* pada penelitian ini memanfaatkan perhitungan *Fuzzy C-Means* serta metode *Elbow* untuk metode penentuan kluster terbaik serta *software* RStudio sebagai aplikasi pengolahan data. Hasil *clustering* dalam penelitian ini 3 kluster adalah kluster terbaik, dibuktikan selisih *SSE* (*Sum of Square Error*) yaitu 1.253, dimana kluster 1 mencakup 15 wilayah, kluster 2 mencakup 12 wilayah dan kluster 3 mencakup 7 wilayah.

**Keywords:** *Clustering, Elbow, Fuzzy C-Means, RStudio, SSE.*

**Abstract**

*Indonesia is a country based on democracy. As for Abraham Lincoln's statement about democracy, said that the perpetrators of the government are the people. Related parties measure the progress of democracy with the measurement model of the Indonesian Democracy Index. The purpose of the Indonesian Democracy index is to read the achievements and develop planning programs to build politics. Clustering is a process of dividing data from a set into several groups where one data with other data if it has a more dominant similarity will form a group, and other data will separate themselves by forming different groups in different groups. This research was conducted to find out new information and knowledge regarding the causes of the decline in scores based on the 2019 Indonesian Democracy Index data and to find out the best clusters in the regrouping of provinces in Indonesia based on the 2019 Indonesian Democracy Index data. Clustering in this study uses the Fuzzy C-Means algorithm and the Elbow method as an optimization method and RStudio software as a data processing application. The results of clustering in this study 3 clusters are the best clusters, as evidenced by the difference in SSE (Sum of Square Error) of 1,253, where cluster 1 covers 15 regions, cluster 2 covers 12 regions and cluster 3 covers 7 regions.*

**Keywords:** *Clustering, Elbow, Fuzzy C-Means, RStudio, SSE.*

## 1. PENDAHULUAN

Indonesia merupakan sebuah negara yang berasaskan demokrasi. Menurut pernyataan Abraham Lincoln mengenai demokrasi adalah *rule of law* dari rakyat, oleh rakyat dan untuk rakyat (Hidayat & Azmi, 2018). Pihak Program Pembangunan Perserikatan Bangsa-Bangsa (UNDP) serta Badan Perencanaan Pembangunan Nasional (BAPPENAS) melakukan pengukuran kemajuan demokrasi dengan model pengukuran Indeks Demokrasi Indonesia (IDI). Fungsinya adalah untuk membaca pencapaian dan menyusun program-program perencanaan membangun politik pada sebuah demokrasi (Ibrahim, 2017).

Menurut Rahayu (2021) Indeks Demokrasi Indonesia lebih rendah dari Timor Leste, tahun 2019 Indonesia menduduki peringkat 64 dunia dengan nilai 6,3 dengan posisi Indonesia lebih rendah dari Malaysia, Timor Leste dan Filipina. Pada tahun 2020 menduduki peringkat 107 dengan nilai 71,8 pertumbuhannya kurang lebih 0,03% dari sebelumnya dengan posisi Indonesia berada di bawah Malaysia dan Thailand. Namun belum diketahui pasti penyebab dari menurunnya skor Indeks Demokrasi Indonesia tahun 2019. Maka di perlukan nya sebuah tindakan berupa pengelompokan data IDI tahun 2019 dari masing-masing provinsi di Indonesia untuk mengetahui informasi dan pengetahuan baru mengenai penyebab terjadinya penurunan skor pada tahun 2019 dan mengetahui karakteristik dari masing-masing provinsi di Indonesia. Pada studi kasus “Analisis kluster hierarki untuk mengelompokkan kembali provinsi-provinsi di Indonesia berdasarkan indeks demokrasi Indonesia 2016” yang dilakukan oleh Ghaisani & Dkk (2019), menggunakan data IDI tahun 2016 dan metode linkage rata-rata, centroid, linkage lengkap, linkage tunggal dan ward, namun pada percobaan tersebut belum ada pengukuran cluster optimum untuk menentukan cluster terbaik. Maka berdasarkan masalah yang di angkat dan juga berdasarkan penelitian terdahulu, maka dilakukan sebuah penelitian terhadap data IDI tahun 2019 pada 34 wilayah di Indonesia yang sudah termasuk provinsi baru Kalimantan Utara. Perhitungan yang digunakan

merupakan perhitungan *Fuzzy C-Means* serta metode *Elbow* sebagai metode penentuan kluster terbaik dan *Software RStudio* untuk pengolahan data. Sehingga, dilakukan penelitian dengan tema “Penerapan *Fuzzy C-Means* dan *elbow* sebagai pengelompokan daerah di Indonesia bersumber pada Indeks Demokrasi Indonesia”.

## 2. TINJAUAN PUSTAKA

### A. Provinsi di Indonesia

Indonesia merupakan negara kepulauan berada antara benua Asia dan Australia. Sesuai Pepres Nomor 6 Tahun 2017 tentang Peraturan Pulau Kecil Terluar, secara geografis, Indonesia mempunyai batas - batas negara dan samudera. Indonesia terdiri dari 34 provinsi dan 13.466 pulau (Suhariyanto, 2020).

### B. Indeks Demokrasi Indonesia

Indeks demokrasi Indonesia merupakan alat ukur sebagai petunjuk tingkat perkembangan demokrasi di Indonesia dengan mengukur 3 aspek yaitu: Kebebasan Sipil, Hak-Hak Politik, dan Lembaga Demokrasi (Suhariyanto, 2018).

### C. Data Mining

*Data mining* merupakan proses mencari informasi pada sekumpulan data menggunakan metode deskripsi, estimasi, prediksi, klasifikasi, pengelompokan (*clustering*) dan asosiasi (Muningsih, 2017).

### D. Clustering

*Clustering* merupakan proses pengelompokan data dimana datanya memiliki kemiripan yang dominan pada kelompok satu dari pada kemiripan data dengan kelompok lainnya (Rofiqo & Dkk, 2018).

### E. Fuzzy C-Means

Algoritma *Fuzzy C-Means* adalah suatu teknik untuk menentukan kluster yang optimal dengan menggunakan derajat keanggotaan sebagai dasar untuk mendefinisikan sebuah vektor yang masuk ke dalam tertentu tertentu. Selain itu *Fuzzy C-Means* dapat menetapkan

kluster bervariasi dan terukur (Butarbutar & Dkk, 2016).

#### F. Metode Elbow

*Elbow* ialah cara untuk mendapatkan pengetahuan tentang kuantitas kluster optimal menggunakan selisih terbesar nilai SSE setiap kluster membentuk tikungan pada sebuah pusat titik *cluster* nya (Winarta & Kurniawan, 2021).

#### G. Rstudio

RStudio merupakan aplikasi tambahan untuk mengolah data, *software* ini mempunyai 2 model, yaitu: *open source* (gratis) dan *versi komersial* (berbayar), selain itu terdapat bentuk aplikasi desktop maupun RStudio *Server*.

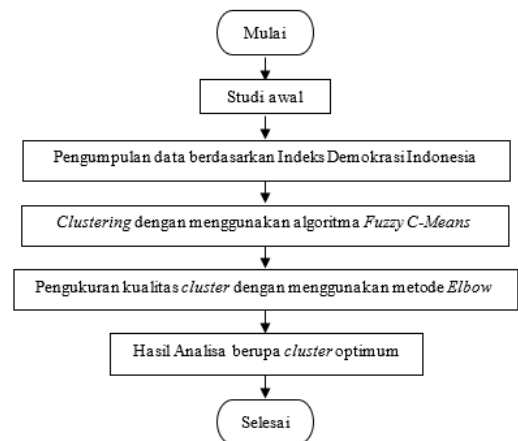
### 3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dan metode eksperimen/pengujian. Menurut Rahmayani & Effendi (2019) metode pendekatan kuantitatif adalah metode yang digunakan untuk memeriksa suatu himpunan dengan menganalisis data yang bersifat statistik. Metode pengujian adalah cara yang dimanfaatkan dalam menemukan efektivitas khusus yang ditujukan untuk mempelajari kemungkinan hubungan sebab-akibat dengan menguji suatu variabel (Arifin, 2018).

*Clustering* atau pengelompokan adalah cara memecah informasi ke dalam golongan dimana informasinya memiliki kemiripan yang lebih dominan dalam suatu kelompok dari pada kemiripan informasi pada beda golongan (Rofiqo & Dkk, 2018). *Clustering* dalam percobaan ini menggunakan perhitungan *Fuzzy C-Means*. Menurut Butarbutar & Dkk (2016) *Fuzzy C-Means* kerap dipakai sebagai pengenalan pola. *Fuzzy C-Means* pada kelompok data dilakukan sesuai dengan derajat keanggotaan yang mana yang berjumlah 0 serta 1.

*Elbow* ditentukan dengan melihat selisih terbesar antara nilai SSE (*Sum of Square Error*) setiap *cluster* membentuk siku pada sebuah titik *cluster* nya (Winarta & Kurniawan, 2021). Menurut Dewi & Pramita (2019), jika *cluster* pertama dan kedua membentuk sudut dan nilai yang paling banyak mengalami penurunan, maka kluster yang paling tepat. RStudio merupakan *software* yang digunakan untuk

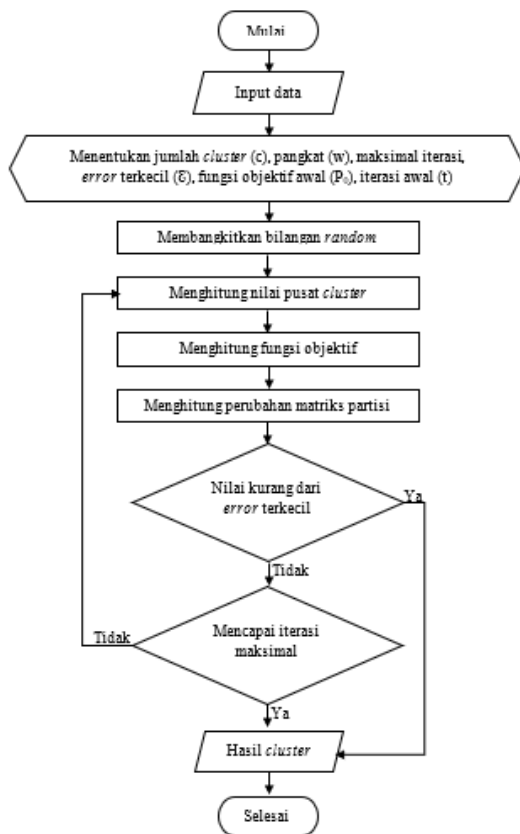
analisis data pengolah informasi dengan menggunakan bahasa R dalam bentuk grafik (Hadiyatullah & Dkk, 2019). Adapun alur pada penelitian ini:



**Gambar 1.** Tahapan Penelitian  
Sumber: (Hadiyatullah & Dkk, 2019)

Tahap awal penelitian ini adalah mencari beberapa literatur terkait dari penelitian sebelumnya sebagai referensi yang dapat mendukung penelitian yang akan dilakukan.

Dataset yang dipakai data sekunder. Data kutip dalam halaman formal BPS. Data yang dipakai ialah data Indeks Demokrasi Indonesia tahun 2019 di 34 provinsi di Indonesia yang mencakup 3 parameter meliputi: AKS, AHP dan ALD.



**Gambar 2.** Flowchart Fuzzy C-Means  
 Sumber: (Butarbutar & Dkk, 2016)

Tahap algoritma *Fuzzy C-Means*:

- 1) Menetapkan data hendak di cluster berupa matriks  $n \times m$ .
- 2) Menetapkan total cluster, *rank*, iterasi maksimal (maksIter), kesalahan minimum, tujuan awal ( $P_0 = 0$ ) serta *iteration* pertama ( $T = 1$ ).
- 3) Menaikkan angka sembarang untuk komponen bagan pemisah pertama U. Mengestimasi nilai tiap ruang dengan rumus:

$$Q_j = \sum_k^c = 1\mu_{ik} \quad (1)$$

Menghitung nilai komponen matriks partisi anggota himpunan U dengan rumus:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \quad (2)$$

- 4) Mengestimasi nilai titik pusat cluster.

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ij})^w \times X_{ij}}{\sum_{i=1}^n (\mu_{ij})^w} \quad (3)$$

Dimana:

$V_{kj}$  = Pusat cluster

$I$  = Iteration

$\mu_{ik}$  = Pergantian matrix partisi

$X_{ij}$  = Attribute

- 5) Mengestimasi fungsi target berulang. Fungsi ini merupakan kondisi iteratif untuk mendapatkan pusat cluster yang tepat.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (4)$$

Dimana:

$P_t$  = Fungsi Objektif

$\sum_{i=1}^n$  = Total data di cluster

$\sum_{i=1}^n$  = Total kalkulasi kluste pertama

- 6) Menghitung pergantian bagan pemisah U:

$$\mu_{ik} = \frac{\left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}}{\sum_{k=1}^c \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}} \quad (5)$$

- 7) Mengecek keadaan berakhir:

- a) Apabila  $(|P_t - P_{t-1}|) / (t > \text{MaksIter})$  berakhir.
- b) Apabila bukan:  $T=T+1$ , kembali pada tahap 4 sampai tujuh.

Metode *Elbow* bertujuan sebagai penentu jumlah cluster terbaik/paling optimum (Winarta & Kurniawan, 2021). Tahapan *Elbow*:

1. Menetapkan nilai pertama cluster (2 cluster).
2. Menaikkan nilai cluster sesuai yang ditentukan.
3. Mengestimasi nilai SSE (*Sum of Square Error*) setiap cluster.

$$SSE = \sum_{k=1}^k \sum_{xi \in Sk} |X_i - C_k|^2 \quad (6)$$

$X_i$  = Nilai attribute informasi ke -  $i$   
 $C_k$  = Angka attribute titik pusat kluster ke -  $i$

4. Menghitung *SSE* hingga kluster yang ditetapkan.
5. Mengamati nilai *SSE* kluster yang turun secara relevan.
6. Menentukan kluster yang membentuk  $t$ .

RStudio merupakan *software* yang dimanfaatkan pada bidang analisis untuk pengolah informasi data dengan bahasa R. RStudio mempunyai dua bentuk, meliputi: *open source/gratis* serta *commercial edition/berbayar*, *software* ini ada bentuk aplikasi desktop maupun bentuk *Server*/yang diakses melalui *google* pada saat komputer terhubung dengan jaringan. RStudio memiliki empat *layout*, diantaranya:

1. *Source/Editor window*, digunakan saat ingin membuat, mengedit serta menyimpan *script R*.
2. *Console window*, digunakan mengeksekusi *command* dari *script R* yang sudah dibuat sebelumnya.
3. *Environment/History window*, *environment* digunakan untuk menampilkan data-data tersimpan di *memory RAM*. Tab *history* digunakan untuk menampilkan daftar *command* yang sudah dijalankan sebelumnya pada *session* yang aktif.
4. Fitur tambahan, digunakan untuk manajemen file, menampilkan *output command* berbentuk *plot*.

#### 4. HASIL DAN PEMBAHASAN

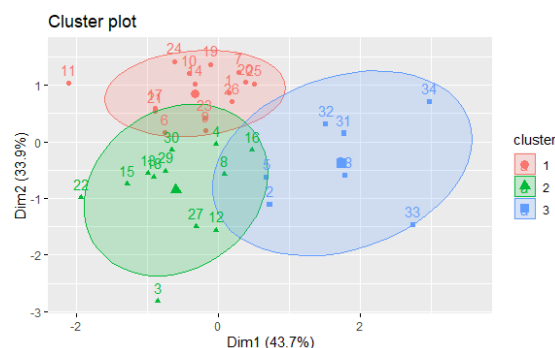
Dataset diolah memakai algoritma *Fuzzy C-Means* serta *Elbow* untuk penentuan kluster terbaik serta *software* RStudio sebagai subkontraktor. Data yang digunakan merupakan data Indeks Demokrasi Indonesia (IDI) pada 34 provinsi di Indonesia tahun 2019 yang mencakup 3 aspek: AKS, AHP serta ALD. Berikut merupakan data sampel IDI tahun 2019 menurut 3 aspek.

**Tabel 1.** IDI Menurut 3 Aspek

Provinsi	AKS	AHP	ALD
Aceh	93,28	65,22	79,08
Sumatera Utara	72,54	61,59	71,12
Sumatera Barat	56,58	61,06	92,21
Riau	85,15	61,68	84,19
Jambi	76,67	64,63	69,21
Sumatera Selatan	80,54	81,95	72,23
Bengkulu	93,98	73,71	68,63
Lampung	78,46	63,70	79,22
Kep. Bangka	84,12	71,71	75,48
Belitung			
Kep. Riau	92,66	78,30	73,10

Sumber: Badan Pusat Statistik

Pengolahan data menggunakan RStudio dan algoritma *Fuzzy C-Means* sebagai algoritma *clustering* dengan skenario 2 hingga 10 kluster menghasilkan 3 kluster optimum dan masing-masing *cluster* mempunyai anggota provinsi.



**Gambar 3.** Plot 3 cluster pada RStudio  
 Sumber: Hasil Perhitungan

Setelah melakukan perhitungan *clustering Fuzzy C-Means*, dilakukan perhitungan *Elbow* pada RStudio untuk menetapkan kluster optimal. Berikut merupakan nilai *Sum of Square Error (SSE)* dari perhitungan metode *Elbow*:

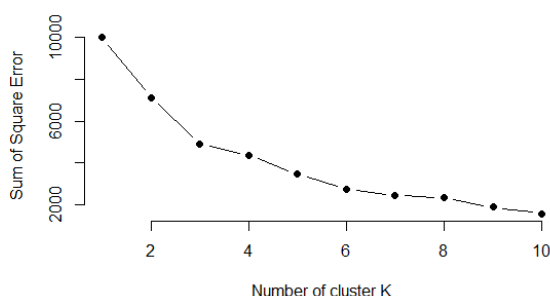
**Tabel 2.** Hasil Metode *Elbow*

C	SSE	Jarak	Keterangan
1	12.558,26	-	-
2	9.560,42	2.998	Selisih C1-C2
3	8.307,58	1.253	Selisih C2-C3
4	6.480,26	1.827	Selisih C3-C4
5	5.583,32	897	Selisih C4-C5
6	4.867,75	716	Selisih C5-C6
7	3.904,09	964	Selisih C6-C7
8	3.548,74	355	Selisih C7-C8



C	SSE	Jarak	Keterangan
9	2.929,63	619	Selisih C8-C9
10	2.747,84	182	Selisih C9-C10

Sumber: Badan Pusat Statistik



**Gambar 4.** Grafik Nilai Metode *Elbow*  
 Sumber: Hasil Perhitungan

Menurut Nainggolan & Dkk (2019) dapat dikategorikan sebagai hasil dari metode *Elbow* atau *cluster* terbaik jika salah satu diantara *cluster* membentuk sudut di sebuah grafik. Apabila kelompok 1 dan kelompok 2 membentuk siku pada titik/nilai memiliki penyusutan terbesar, *cluster* tersebut merupakan *cluster* optimum. Dan menurut Yuan & Yang (2019) dapat dikatakan sebagai *cluster* terbaik apabila nilai SSE yang lebih kecil menunjukkan bahwa setiap *cluster* lebih konvergen. Adapun menurut Saputra & Dkk (2019) metode *Elbow* ditunjukkan ketika grafik *Elbow* memiliki tikungan yang signifikan. Sehingga *cluster* 3 merupakan *cluster* optimum.

Dari proses *clustering* yang sudah dilakukan menunjukkan letak *cluster* optimum terdapat pada *cluster* 3. Untuk melihat ciri-ciri setiap *cluster*, perlu membuat profil *cluster* pada data IDI di setiap masing-masing provinsi yang termasuk pada *cluster* 1, *cluster* dua, serta *cluster* tiga berdasarkan perhitungan yang dibuat sebelumnya. Dan penentuan ciri-ciri setiap *cluster* diambil berdasarkan nilai Tertinggi (*Max*) dan Terendah (*Min*) (Novany & Safii, 2021). Hasil 3 *cluster* yang terdiri dari *cluster* 1, *cluster* 2, dan *cluster* 3 di tampilkan pada tabel berikut.

**Tabel 3.** Anggota Provinsi *Cluster* 1

Provinsi	AKS	AHP	ALD
Aceh	93,28	65,22	79,08
Sumatera Selatan	80,54	81,95	72,23

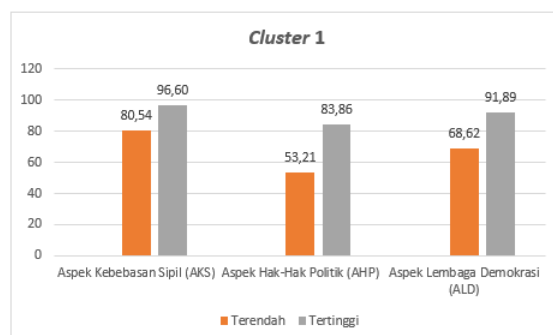
Provinsi	AKS	AHP	ALD
Bengkulu	93,98	73,17	68,63
Kep. Bangka Belitung	84,12	71,71	75,48
Kep. Riau	92,66	78,30	73,10
DKI Jakarta	91,01	83,86	91,89
DI Yogyakarta	92,90	72,51	78,25
Bali	90,60	68,38	90,42
Nusa Tenggara Timur	93,97	78,58	68,62
Kalimantan Barat	92,59	71,45	68,74
Kalimantan Tengah	86,31	78,24	79,32
Kalimantan Timur	88,50	66,27	82,17
Kalimantan Utara	96,60	74,92	80,46
Sulawesi Utara	95,19	63,54	75,82
Sulawesi Tengah	96,19	53,21	91,68

Sumber: Hasil Perhitungan

**Tabel 4.** Karakteristik *Cluster* 1

Kluster 1	AKS	AHP	ALD
Terendah	80,54	53,21	68,62
Tertinggi	96,60	83,86	91,89

Sumber: Hasil Perhitungan



**Gambar 5.** Karakteristik *Cluster* 1  
 Sumber: Hasil Perhitungan

Tabel di atas menunjukkan bahwa setiap *cluster* mempunyai anggota provinsi dan karakteristik masing-masing yang ditunjukkan dari nilai Terendah (*MIN*), dan Tertinggi (*MAX*). Tabel 3 merupakan anggota dari *cluster* 1 yang terdiri dari 15 anggota provinsi. Tabel 4 merupakan karakteristik pada *cluster* 1, yaitu: parameter Aspek Kebebasan Sipil (AKS) mempunyai nilai terendah 80,54 dan nilai tertinggi 96,60. Parameter Aspek Hak-Hak Politik (AHP) mempunyai nilai terendah 53,21 dan nilai tertinggi 83,86. Parameter Aspek Lembaga Demokrasi (ALD) mempunyai nilai terendah 68,62 dan nilai tertinggi 91,89. Dan

karakteristik tersebut juga ditunjukkan dalam bentuk grafik pada gambar 5. Sehingga hasil dari perhitungan karakteristik 3 *cluster* tersebut masing-masing *cluster* dapat di simpulkan bahwa *cluster* 1 merupakan *cluster* tertinggi pada parameter Aspek Kebebasan Sipil (AKS) dengan nilai 96,60. Dan *cluster* 3 merupakan *cluster* terendah pada parameter Aspek Hak-Hak Politik (AHP) dengan nilai 50,31.

## 5. KESIMPULAN DAN SARAN

### A. KESIMPULAN

Penerapan *clustering* memakai perhitungan *Fuzzy C-Means* serta cara *Elbow* berdasarkan data Indeks Demokrasi Indonesia (IDI), 3 *cluster* merupakan kluster terbaik berdasarkan selisih *Sum of Square Error (SSE)* 1.253, dan setiap *cluster* mempunyai anggota provinsi serta karakteristik yang ditunjukkan dari nilai Terendah (*MIN*) dan Tertinggi (*MAX*).

Kluster 1 termasuk 15 provinsi, serta karakteristik pada kluster 1, yaitu: parameter Aspek Kebebasan Sipil (AKS) mempunyai nilai terendah 80,54 dan nilai tertinggi 96,60. Parameter Aspek Hak-Hak Politik (AHP) mempunyai nilai terendah 53,21 dan nilai tertinggi 83,86. Parameter Aspek Lembaga Demokrasi (ALD) mempunyai nilai terendah 68,62 dan nilai tertinggi 91,89. *Cluster* 2 termasuk 12 provinsi, serta karakteristik pada kluster 2, yaitu: parameter Aspek Kebebasan Sipil (AKS) mempunyai nilai terendah 56,58 dan nilai tertinggi 85,15. Parameter Aspek Hak-Hak Politik (AHP) mempunyai nilai terendah 60,29 dan nilai tertinggi 86,52. Parameter Aspek Lembaga Demokrasi (ALD) mempunyai nilai terendah 70,75 dan nilai tertinggi 92,21. *Cluster* 3 terdiri dari 7 anggota provinsi, dan karakteristik pada *cluster* 3, yaitu: parameter Aspek Kebebasan Sipil (AKS) mempunyai nilai terendah 70,35 dan nilai tertinggi 93,08. Parameter Aspek Hak-Hak Politik (AHP) mempunyai nilai terendah 50,31 dan nilai tertinggi 64,86. Parameter Aspek Lembaga Demokrasi (ALD) mempunyai nilai terendah 52,61 dan nilai tertinggi 71,12.

### B. SARAN

Saran dari penelitian yang telah dilakukan adalah sebagai berikut.

1. Dalam proses validasi mencari *cluster* optimum dapat menggunakan alternatif lain seperti *Davies Bouldin Index*, *Silhouette Coefficient*, dan *Calirnski-Harabasz index*.
2. Penggunaan *software* untuk membandingkan hasil *clustering* dapat menggunakan alternatif lain seperti *Orange*, *Weka*, dan *RapidMiner*.

## 6. REFERENSI

- Arifin, Z. (2018). *Metodologi Penelitian Pendidikan Education Research Methodology*. 1–5.
- Butarbutar, & Dkk. (2016). Komparasi Kinerja Algoritma Fuzzy C-Means dan K-Means dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa ( Studi Kasus : SMP Negeri 2 Pematangsiantar ). *Jurnal Riset Informasi & Teknis Informatika*, 1, 46–55.
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix : Jurnal Manajemen Teknologi Dan Informatika*, 9, 102–109.
- Ghaisani, & Dkk. (2019). Analisis Cluster Hirarki Untuk Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Demokrasi Indonesia Tahun 2016. *Prosiding Konferensi Nasional Penelitian Matematika Dan Pembelajarannya (KNPMP) IV*.
- Hadiyatullah, & Dkk. (2019). Algoritma Fuzzy C-Means Dengan Metode Elbow Untuk Mengelompokkan Kepolisian Provinsi Di Indonesia Berdasarkan Jumlah Kasus ( Crime Total , Crime Cleared , Creime Rate , Clearance Rate ). (*JASIE*) *Jurnal Aplikasi Sistem Informasi Dan Elektronika*, 1, 1–18.
- Hidayat, T., & Azmi, K. (2018). *Pembangkaian Berita Dugaan Kecurangan Pilkada Serentak 2017 DKI Jakarta Di Detik.Com Periode 18 Februari 2017” (Analisis Framing Robert n Entmant)*. 1–10.

- Ibrahim. (2017). Menakar kedalaman pengukuran demokrasi model Indeks Demokrasi Indonesia ( IDI ): Beberapa catatan substantif dari Kepulauan Bangka Belitung. *Jurnal Masyarakat, Kebudayaan Dan Politik*, 131–149.
- Muningsih, E. (2017). Optimasi Jumlah Custer K-Means Dengan Metode Elbow Untuk Pemetaan Pelanggan. *Prosiding Seminar Nasional ELINVO*, 105–114.
- Nainggolan, & Dkk. (2019). Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method. *Journal of Physics: Conference Series*, 1361(1), 1–6. <https://doi.org/10.1088/1742-6596/1361/1/012015>
- Novany, A. A., & Safii, M. (2021). Penerapan Data Mining Dalam Mengelompokkan Produksi Daging Sapi di Pulau Sumatera Menggunakan Algoritma K-Means Clustering. *Seminar Nasional Informatika (SENATIKA)*, 237–243.
- Rahayu, L. S. (2021). *Ma'ruf Amin Ungkap Indeks Demokrasi RI di Bawah Timor Leste*. News.Detik.Com. <https://news.detik.com/berita/d-5546931/maruf-amin-ungkap-indeks-demokrasi-ri-di-bawah-timor-leste>
- Rahmayani, S. R., & Effendi, K. N. S. (2019). Kemampuan Komunikasi Matematis Siswa SMP Pada Materi Himpunan. *JUDIKA (JURNAL PENDIDIKAN UNSIKA)*, 7(1), 10–18. <http://journal.unsika.ac.id/index.php/judika>
- Rofiqo, & Dkk. (2018). Penerapan Clustering Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 2, 216–223.
- Saputra, & Dkk. (2019). Effect of Distance Metrics in Determining K-Value in KMeans Clustering Using Elbow and Silhouette Method. *Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN)*, 172, 341–346.
- Suhariyanto. (2018). *Berita Resmi Statistik Indeks Demokrasi Indonesia (IDI) 2017*. Badan Pusat Statistik.
- Suhariyanto. (2020). *Statistik Indonesia 2020*. Badan Pusat Statistik.
- Winarta, A., & Kurniawan, W. J. (2021). Optimasi Cluster K-Means Menggunakan Metode Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python. *Jurnal Teknik Informatika Kaputama (JTIK)*, 5, 113–119.
- Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *Multidisclipinary Scientific Journal*, 2(2), 227–235. <https://doi.org/10.3390/j2020016>