

**PERBANDINGAN METODE K-MEANS DAN FUZZY C-MEANS TERHADAP
INDIKATOR KEMISKINAN NTT 2021
(Studi Kasus : Indikator Tingkat Kemiskinan di NTT tahun 2021)**

Alfriyani R D Klau¹, Kris Suryowati^{2*}

^{1,2} Jurusan Statistika, FST, Institut Sains & Teknologi AKPRIND Yogyakarta

E-mail: alfrynahak@gmail.com, suryowati@akprind.ac.id

*corresponding author

Abstract: Poverty is a condition where there is an inability to meet basic needs such as food, clothing, shelter, education and health. The purpose of this analysis is to determine the characteristics of the poverty rate in East Nusa Tenggara Province, to determine the factors of the poverty rate in East Nusa Tenggara Province and to determine the best model analysis, to know the results of grouping poverty levels based on influencing factors using the K method. - Means, find out the results of grouping poverty levels based on influencing factors using the K-Means method, find out which method is the best between the K-Means and Fuzzy C-Means methods. Some of the methods in clustering include the K-Means method and Fuzzy C-means . The results of the analysis can be concluded that the best method in the 2021 Poverty case study in East Nusa Tenggara is the K-Means method with Davies Bouldin values Index and R-Squared. of 1,615 with a total of 2 clusters. The two clusters are grouped into areas that have a high and low impact on the poverty rate in the Province of East Nusa Tenggara in 2021. The areas that have a high effect on the poverty rate in the Province of East Nusa Tenggara in 2021 are in cluster 2.

Keywords: Poverty rate, K-Means, Fuzzy C-Means, Davies Bouldin Index, R-Square

Abstrak: Kemiskinan merupakan keadaan yang terjadi ketidak mampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan. Tujuan dalam analisis ini adalah untuk mengetahui karakteristik tingkat kemiskinan di Provinsi Nusa Tenggara Timur, untuk mengetahui faktor-faktor dari tingkat kemiskinan di Provinsi Nusa Tenggara Timur dan menentukan analisis model terbaik, mengetahui hasil pengelompokan tingkat kemiskinan berdasarkan factor-faktor yang mempengaruhi dengan menggunakan metode K-Means, mengetahui hasil pengelompokan tingkat kemiskinan berdasarkan factor-faktor yang mempengaruhi dengan menggunakan metode K-Means, mengetahui metode manakah yang terbaik antara metode K-Means dan Fuzzy C-Means. Beberapa metode yang ada dalam clustering antara lain metode K-Means dan Fuzzy C-means. Hasil analisis dapat disimpulkan bahwa metode terbaik dalam studi kasus Kemiskinan di Nusa Tenggara Timur tahun 2021 adalah metode *K-Means* dengan nilai nilai *Davies Bouldin Index* dan *R-Squared*. sebesar 1.615 dengan jumlah *cluster* sebanyak 2 *cluster*. Dari dua *cluster* tersebut dikelompokkan menjadi wilayah yang tinggi dan rendah mempengaruhi tingkat kemiskinan di Provinsi Nusa Tenggara Timur Tahun 2021. Daerah yang tinggi mempengaruhi tingkat kemiskinan di Provinsi Nusa Tenggara Timur Tahun 2021 terdapat pada cluster 2.

Kata kunci : Tingkat kemiskinan, K-Means, Fuzzy C-Means, Davies Bouldin Index, R-Squared.

1. PENDAHULUAN

Kemiskinan merupakan keadaan yang terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan. Kemiskinan dapat disebabkan oleh kelangkaan alat pemenuh kebutuhan dasar, ataupun sulitnya akses terhadap pendidikan dan pekerjaan. Kemiskinan terjadi karena kemampuan masyarakat pelaku

ekonomi tidak sama, sehingga terdapat masyarakat yang tidak dapat ikut serta dalam proses pembangunan atau menikmati hasil pembangunan (Soegijoko, 2001). Rumusa masalah dalam analisis ini Bagaimana hasil pengelompokan kedua metode dan dari kedua metode ini, metode manakah yang terbaik dalam analisis data Tingkat Kemiskinan. Tujuan dari penelitian ini adalah untuk mengetahui hasil dari rumusan masalah

2. METODE

1. Desain penelitian

Desain penelitian yang digunakan adalah pendekatan kuantitatif deskriptif yang dimana pendekatan disajikan dalam bentuk angka atau numerik dan interpretasi hasil dilakukan dalam bentuk deskripsi.

2. Objek penelitian

Objek penelitian pada penelitian ini adalah 22 kabupaten/kota yang berada di provinsi Nusa Tenggara Timur. Adapun daftar Kabupaten/ kota tersebut disajikan dalam bentuk Tabel 2.1 berikut:

Tabel 2.1 (Daftar Nama Kabupaten/ Kota di NTT)

| No | Kota / Kabupaten | No | Kota / Kabupaten |
|----|----------------------|----|------------------|
| 1 | Sumba Barat | 12 | Ngada |
| 2 | Sumba Timur | 13 | Manggarai |
| 3 | Kupang | 14 | Rote Ndao |
| 4 | Timor Tengah Selatan | 15 | Manggarai Barat |
| 5 | Timor Tengah Utara | 16 | Sumba Tengah |
| 6 | Belu | 17 | Sumba Barat Daya |
| 7 | Alor | 18 | Nagekeo |
| 8 | Lembata | 19 | Mangarai Timur |
| 9 | Flores Timur | 20 | Sabu Raijua |
| 10 | Sikka | 21 | Malaka |
| 11 | Ende | 22 | Kota Kupang |

3. Sumber Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan cara publikasi digital website resmi Badan Pusat Statistika Provinsi Nusa Tenggara Timur (<https://ntt.bps.go.id>).

4. Variable Penelitian

Pada penelitian ini data yang digunakan adalah jenis data sekunder berupa laju pertumbuhan penduduk, presentase angka buta huruf penduduk berumur 15 tahun ke atas, laju pertumbuhan PDRM atas dasar harga konstan, persentase jumlah penduduk miskin, laju pertumbuhan penduduk, Angka Partisipasi Sekolah (APS) Umur 5-6Tahun, presentase yang mempunyai keluhan Kesehatan dan mengakibatkan terganggunya kegiatan dalam 1 bulan. di Provinsi Nusa Tenggara Timur (NTT) menurut kabupaten/kota (persen) tahun 2021.

5. Metode Analisis

Metode yang digunakan untuk menganalisis data adalah Algoritma *K-Means* dan *Fuzzy C Means*. Pengolahan data dalam penelitian ini menggunakan *software Microsoft Excel* dan *R studio*. Berikut langkah-langkah yang perlu dilakukan:

- a. Menyiapkan data Jumlah Kemiskinan tahun 2021.

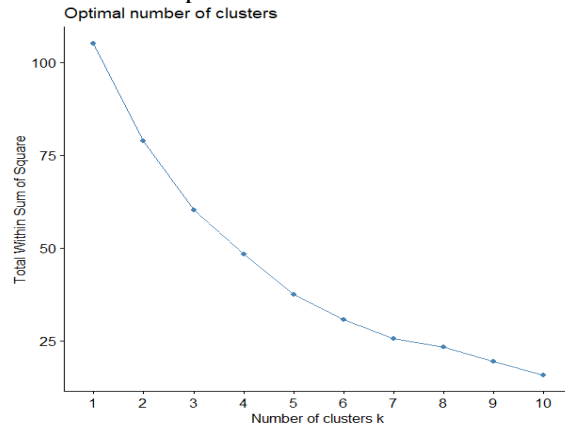
Melakukan analisis deskriptif menggunakan *software Microsoft Excel*

- b. Melakukan uji multikolinearitas untuk mengetahui atau melihat korelasi antar variabel dan sebagai asumsi yang harus dipenuhi dalam proses *clustering*
- c. Melakukan pengelompokan kabupaten/kota di Nusa Tenggara Timur menggunakan metode Algoritma *K-Means Clustering*. Langkah-langkah analisis data menggunakan metode *K-Means* adalah sebagai berikut:
 - i. Melakukan standarisasi data menggunakan *software Microsoft Excel*.
 - ii. Menentukan jumlah *cluster* yang optimal menggunakan metode *Elbow*.
 - iii. Menghitung jarak antar titik *centroid* dengan titik tiap objek.
 - iv. Pengelompokan objek untuk menentukan anggota *cluster* adalah dengan memperhitungkan jarak minimum objek.
 - v. Kembali ke langkah b jika anggota *cluster* berpindah. Lakukan perulangan hingga nilai *centroid* yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke *cluster* lain.
- d. Melakukan pengelompokan kabupaten/kota di Nusa Tenggara Timur dengan menggunakan metode Algoritma *Fuzzy C-means*. Langkah-langkah analisis data menggunakan metode *Fuzzy Fuzzy C-Means* adalah sebagai berikut:
 - a. Memasukkan data berupa matriks berukuran $n \times m$ yang memiliki fungsi untuk menentukan jumlah data dan jumlah atribut atau variabel setiap data yang akan digunakan.
 - b. Menentukan:
 - a) Jumlah cluster = $c (\geq 2)$;
 - b) Bobot pangkat = $w (> 1)$;
 - c) Maksimum iterasi = MaxIter
 - d) Error terkecil yang diharapkan = ξ
 - e) Fungsi objektif awal = $P_0 = 0$
 - f) Iterasi awal = $t = 1 = 1000$
 - g) Membangkitkan bilangan random untuk mengisi elemen-elemen matriks partisi awal U atau matriks kekhasan relatif dengan rentang nilai 0 sampai 1
 - c. Menghitung pusat cluster ke- k
 - d. Menghitung fungsi objektif pada iterasi ke = t
 - e. Menghitung perubahan matriks partisi
 - f. Mengecek kondisi berhenti
 - g. Membandingkan metode yang terbaik antara algoritma *K-Means* dan *Fuzzy C-Means*
 - h. Menarik kesimpulan

3. HASIL DAN PEMBAHASAN

1. Analisis pengelompokan metode K-Means terhadap tingkat kemiskinan di Provinsi Nusa Tenggara Timur tahun 2021

- a. Menentukan Jumlah *Cluster* Optimal



Gambar 1.1

Dapat dilihat bahwa grafik yang pergerakannya membentuk siku atau menurun secara drastis adalah grafik pada titik angka 2. Hal ini berarti bahwa nilai k optimum pada metode *Elbow* adalah 2.

- b. Menentukan Pusat *Cluster*

Table 2.2

| <i>Centroid</i> | X_1 | X_2 | X_3 | X_4 | X_5 |
|-----------------|--------|--------|--------|--------|--------|
| 1 | 0.296 | -0.186 | -0.954 | 0.588 | -0.706 |
| 2 | -0.205 | 0.129 | 0.661 | -0.407 | 0.488 |

- c. Jarak antar titik pusat *cluster* dengan tiap objek

- Jarak setiap kabupaten/kota dengan *centroid* pertama (c_1)

$$d(1, c_1) = \sqrt{(1.611 - 0.296)^2 + (1.867 - (-0.186))^2 + (-1.161 - (-0.954))^2 + (-0.692 - 0.588)^2 + (-1.517 - (-0.706))^2}$$

- d. Pengelompokan anggota *cluster* berdasarkan Jarak terdekat

Berdasarkan hasil penentuan jumlah *cluster* optimal yaitu $k = 2$ dengan menggunakan data indikator variabel tingkat kemiskinan di Provinsi Nusa Tenggara Timur, maka didapat jumlah setiap *cluster* adalah *cluster* 1 memiliki 9 anggota dan *cluster* 2 memiliki 13 anggota.

2. Hasil pengelompokan metode fuzzy c-means terhadap indikator tingkat kemiskinan di Provinsi Nusa Tenggara Timur tahun 2021

Table 2.3

| | |
|------------------|--|
| <i>Cluster</i> 1 | Kabupaten/Kota :1,2,6,8,9,12,13,15,20,22 |
|------------------|--|

| | |
|------------------|---|
| <i>Cluster 2</i> | Kabupaten/Kota :3,4,5,7,10,11,14,16,17,18,19,21 |
|------------------|---|

Pada hasil output di atas dapat dilihat bahwa pada *cluster 1* terdiri dari 10 anggota yaitu Kabupaten/Kota 1,2,6,8,9,12,13,15,20,22. Pada *cluster 2* terdapat 12 anggota *cluster* yang terdiri dari Kabupaten/Kota 3,4,5,7,10,11,14,16,17,18,19,21.

3. *Metode terbaik antara k-means dan fuzzy c-means dalam pengelompokan indikator tingkat kemiskinan kabupaten di Provinsi Nusa Tenggara Timur*
 - a. Penentuan metode terbaik.

Tabel 2.4

| | Metode | |
|-----------------------------|----------------|----------------------|
| | | |
| | <i>k-means</i> | <i>fuzzy k means</i> |
| <i>Davies-Bouldin index</i> | 1.615 | 1.829 |
| <i>R-Squared</i> | 27.2% | 14.96% |

Kriteria nilai yang digunakan dalam menentukan metode terbaik adalah dengan menggunakan nilai *Davies Bouldin Index* dan *R- Squared*. Berdasarkan Tabel 4.15 dapat diketahui metode *Davies Bouldin Index* untuk penentuan metode terbaik adalah metode *K-Means* dengan nilai sebesar 1,615 dan *R-Squared* sebesar 27.2% dengan jumlah *cluster* sebanyak 2 *cluster*. Sehingga dapat disimpulkan bahwa metode terbaik dalam pengelompokan kabupaten/kota di provinsi NTT berdasarkan indikator variabel tingkat kemiskinan adalah metode *K-Means Clustering* dengan banyaknya *cluster 2*.

- b. Jumlah Cluster

Table 2.5

| <i>Cluster</i> \Variabel | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ |
|--------------------------|----------------|----------------|----------------|----------------|----------------|
| C ₁ | 1.07 | 4.95 | 1.15 | 38.2 | 11.6 |
| C ₂ | 0.86 | 6.12 | 2.09 | 27.6 | 16.4 |

Dapat diketahui pengelompokan dari masing-masing *cluster* yang dikategorikan berdasarkan nilai rata-rata setiap variabel pada *cluster* yang terbentuk. Pada *cluster 1* dan *cluster 2* nilai rata-rata yang kecil merupakan kelompok wilayah rendah sedangkan nilai rata-rata yang lebih besar merupakan kelompok wilayah yang tinggi.

4. KESIMPULAN

Kriteria nilai yang digunakan dalam menentukan metode terbaik adalah dengan menggunakan nilai *Davies Bouldin Index* dan *R- Squared*. Berdasarkan Tabel 4.15 dapat diketahui metode *Davies Bouldin Index* untuk penentuan metode terbaik adalah metode *K-Means* dengan nilai sebesar 1,615 dan *R-Squared* sebesar 27.2% dengan jumlah *cluster* sebanyak 2 *cluster*. Sehingga dapat disimpulkan bahwa metode terbaik dalam pengelompokan kabupaten/kota di provinsi NTT berdasarkan indikator variabel tingkat kemiskinan adalah metode *K-Means Clustering* dengan banyaknya *cluster 2*.

5. DAFTAR PUSTAKA

- Afifah N, 2016, Pengklasteran Lahan Sawah Di Indonesia Sebagai Evaluasi Ketersediaan Produksi Pangan Menggunakan *Fuzzy C-Means*, *Jurnal Matematika "MANTIK" Volume 2, Nomor 1, Oktober 2016, Universitas Islam Negeri Sunan Ampel Surabaya, Surabaya.*
- Alwi W, ddk, 2018, Analisis Klaster Untuk Pengelompokan Kabupaten/Kota Di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat, *Jurnal MSA Volume 6, Nomor 1, Januari – Juni 2018, Makasar.*
- Asfi M, 2008, Pelabelan Otomatis Citra Menggunakan Fuzzy C-Means Untuk Sistem Temu Kembali Citra, *Volume 2, Bogor, Institut Pertanian Bogor.*
- Anwar H, 2014, Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Index Pembangunan Manusia Tahun 2014, *Jurnal Ilmu Matematika dan Terapan, Volume 11, Nomor 2, Desember 2019, Maluku.*
- Bahar, 2011, Penentuan Jurusan Sekolah Menengah Atas Dengan Algoritma Fuzzy C-Means, *Semarang, Universitas Dian Nuswantoro.*