

ANALISA CLUSTER MENGGUNAKAN K-MEANS DAN FUZZY C-MEANS DALAM PENGELOMPOKAN PROVINSI MENURUT DATA INTESITAS BENCANA ALAM DI INDONESIA TAHUN 2017-2021**Hajar Firdaus**Program Studi Matematika, FMIPA, Universitas Negeri Surabaya, Surabaya, Indonesia
e-mail: hajar.18010@mhs.unesa.ac.id**A'yunin Sofro**Program Studi Matematika, FMIPA, Universitas Negeri Surabaya, Surabaya, Indonesia
Penulis Korespondensi: ayuninsofro@unesa.ac.id**Abstrak**

Bencana merupakan sebuah kejadian yang serius dimana terjadi sebuah kekacauan yang dapat disebabkan oleh alam, manusia, teknologi yang berbahaya, serta berbagai faktor lain yang mempengaruhi keadaan normal suatu wilayah ataupun komunitas. Indonesia merupakan salah satu negara dengan jumlah intensitas bencana alam terbanyak di dunia setelah Amerika Serikat. Pada tahun 2020, Indonesia tercatat memiliki jumlah intensitas bencana alam sebesar 4.977 kasus. Tindakan penanggulangan untuk menangani kekacauan dan juga kerugian yang terjadi dikarenakan bencana alam merupakan salah satu tanggung jawab dari pemerintah di Indonesia. Tindakan penanggulangan bencana alam tersebut bertujuan untuk meminimalisir terjadinya korban jiwa ataupun kerugian lainnya yang disebabkan oleh bencana alam tersebut, sehingga perlu adanya tindakan untuk mencegah banyaknya kerugian yang terjadi. Salah satunya yaitu dengan menentukan karakteristik dari intensitas bencana alam tersebut dan mengelompokkan provinsi-provinsi di Indonesia berdasarkan karakteristik yang dimiliki. Pada artikel ini akan membahas pengelompokan provinsi di Indonesia dengan menggunakan metode *K-means clustering* dan *Fuzzy C-means clustering*. Hasil yang diperoleh dari analisis *clustering* tersebut menunjukkan bahwa dengan menggunakan *Fuzzy C-means clustering* akan memiliki akurasi yang lebih baik dibandingkan dengan *K-means clustering*. Hal tersebut ditunjukkan dengan rata-rata *Silhouette Coefficient Fuzzy C-means clustering* yang lebih besar dibandingkan dengan *K-means clustering* yaitu sebesar 0.8403194 yang termasuk dalam kategori pengelompokan yang kuat.

Kata Kunci: Bencana Alam, Indonesia, K-means, *Fuzzy C-means*.**Abstract**

Disaster is a serious event where there is chaos that can be caused by nature, humans, dangerous technology, and various other factors that affect the normal circumstances of a region or community. Indonesia is one of the countries with the highest number of natural disaster intensity in the world after the United States. In 2020, Indonesia recorded a total natural disaster intensity of 4.977 cases. Countermeasures to deal with chaos and also losses caused by natural disasters are one of the responsibilities of the government in Indonesia. The natural disaster management action aims to minimize the occurrence of fatalities or other losses caused by natural disasters, so there needs to be action to prevent the many losses that occur. One of them is by determining the characteristics of the intensity of the natural disaster and grouping the provinces in Indonesia based on the characteristics they have. In this article will discuss the grouping of provinces in Indonesia using the method of *K-means clustering* and *Fuzzy C-means clustering*. Results obtained from the clustering analysis showed that using *Fuzzy C-means clustering* would have better accuracy compared to *K-means clustering*. This is indicated by the average *Fuzzy Silhouette Coefficient C-means clustering* which is greater than *K-means clustering* which is 0.8403194 which falls into the category of strong grouping.

Keywords: Natural Disaster, Indonesia, K-means, *Fuzzy C-means*

PENDAHULUAN

Bencana merupakan sebuah kejadian yang serius dimana terjadi sebuah kekacauan yang dapat disebabkan oleh alam, manusia, teknologi yang berbahaya, serta berbagai faktor lain yang mempengaruhi keadaan normal suatu wilayah ataupun komunitas (IFRC, 2021; WHO/EHA, 2002). Kekacauan tersebut dapat mengakibatkan terjadinya berbagai kerugian baik dari segi kemanusiaan, ekonomi, material, dan lingkungan yang menjadi kacau balau dan rusak (IFRC, 2021).

Bencana alam merupakan salah satu bencana yang diakibatkan oleh faktor-faktor yang terjadi pada alam antara lain yaitu faktor geologi, hidrologi, meteorologi, klimatologi, biologi, dan faktor yang diakibatkan oleh benda-benda yang ada di luar angkasa (Chaudhary & Piracha, 2021). Contoh bencana alam yang diakibatkan oleh faktor geologi adalah gempa bumi, letusan gunung api, dan pergerakan lempeng bumi, sedangkan bencana alam yang disebabkan oleh faktor hidrologi adalah banjir, tanah longsor dan aktivitas dari gelombang air. Untuk bencana alam yang diakibatkan oleh faktor meteorologi adalah badai konvektif, badai ekstratropis, suhu ekstrim dan kabut, sedangkan untuk bencana alam yang disebabkan oleh faktor klimatologi adalah kekeringan, ledakan danau glasial, kebakaran hutan. Untuk bencana alam yang disebabkan oleh faktor biologi adalah kejadian binatang berbahaya, penyakit, dan infeksi serangga, sedangkan bencana alam yang disebabkan oleh benda-benda yang ada di luar angkasa adalah benturan pada atmosfer bumi, suhu di luar angkasa dan di atmosfer (Glossary, 2014). Terdapat beberapa dampak negatif dari bencana alam tersebut antara lain yaitu kelangkaan bahan pangan, terjadinya trauma pasca bencana, menurunnya daya tahan tubuh, terjadinya migrasi secara besar-besaran, masalah finansial dan ekonomi, serta timbul perasaan khawatir akan kehidupan selanjutnya setelah terjadinya bencana alam di wilayahnya (Chaudhary & Piracha, 2021).

Menurut (UCLouvain dkk., 2019), Indonesia merupakan salah satu negara dengan jumlah intensitas bencana alam terbanyak di dunia setelah Amerika Serikat. Pada tahun 2020, Indonesia tercatat memiliki jumlah intensitas bencana alam sebesar 4.977 kasus, dimana jumlah tersebut merupakan jumlah terbanyak kasus bencana alam yang pernah

terjadi di Indonesia (BNPB, 2021a). Banyaknya kejadian bencana alam di Indonesia dikarenakan letak geografis Indonesia yang terletak pada *ring of fire* di jalur gugus gunung api yaitu sekitar 28.91 GW dari energi panas bumi yang tersebar pada 312 lokasi di berbagai pulau di Indonesia antara lain yaitu Jawa, Sulawesi, Sumatra, Bali, Nusa Tenggara dan Sulawesi (Pambudi, 2018). Bentang alam Indonesia juga terletak pada pertemuan lempeng tektonik yang menyebabkan banyaknya kejadian gunung Meletus dan gempa bumi (Verstappen, 2010). Letak geografis Indonesia juga terletak pada batas antara 3 lempeng utama yaitu Eurasia, India-Australia, dan Laut Pasifik-Philippine (Hall, 2009). Pergerakan dari lempeng tektonik India-Australia tersebut menyebabkan terbentuknya selimut salju pada puncak tertinggi suatu gunung, selimut salju tersebut juga menutupi daerah barat serta timur dari pulau tersebut, dan menyebabkan dataran aluvial yang panas dan lembab di daerah sekitaran pantai (Johnson, 2004). Dikarenakan pergerakan lempeng tektonik India-Australia tersebut juga menyebabkan banyak sekali terjadi bencana alam tsunami dan gempa bumi di Indonesia, khususnya di Pulau Papua bagian barat dan utara (Hillis & Müller, 2003).

Pada tahun 2021 tepatnya pada tanggal 04 Desember, terjadi bencana alam yaitu meletusnya Gunung Semeru yang terjadi pukul 14.50 WIB dan berlokasi di Kabupaten Lumajang, Jawa Timur. Bencana Alam tersebut mengakibatkan banyak sekali korban jiwa, diantaranya yaitu terdapat 34 orang yang meninggal, 169 orang yang terluka, 5.025 orang yang menderitanya, dan 4.350 orang yang menggungsi (BNPB, 2021b). Kerugian yang terjadi akibat bencana alam tersebut tidak hanya berasal dari segi manusia saja, melainkan juga dari segi material yaitu mengakibatkan 2.970 unit rumah yang terdampak akibat erupsi Gunung Semeru tersebut, juga diketahui bahwa terdapat 38 unit Pendidikan yang mengalami kerusakan dan salah satu jembatan yang berada di Kabupaten Lumajang yang bernama Gladak Perak juga putus (Yanuarto, 2021). Bencana alam tersebut membuat masyarakat di sekitaran gunung tersebut menjadi resah dikarenakan akan adanya bencana alam susulan yang akan terjadi di daerah tersebut sedangkan rumah milik para warga tersebut telah hancur dan rusak dikarenakan letusan Gunung Semeru, sehingga banyak sekali warga yang memutuskan untuk menggungsi ke tempat yang lebih

aman ataupun berpindah ke daerah yang tidak rawan terjadi bencana alam (Masyekh & Faradilla, 2021). Adapun beberapa bencana alam yang kerap terjadi di Indonesia yaitu banjir, tanah longsor, kekeringan, kebakaran hutan, gempa bumi dan letusan gunung api.

Analisa *clustering* adalah teknik analisis *unsupervised learning* yang dilakukan dengan tujuan mengelompokkan objek berdasarkan analisis statistik dari setiap objek (Batani dkk., 2017). Salah satu analisa *clustering* paling populer yaitu *partitioning clustering* (Omran dkk., 2007). *Partitioning clustering* adalah sebuah metode *clustering* yang digunakan untuk mengklasifikasi observasi pada *dataset* menjadi beberapa kelompok berdasarkan kemiripan datanya (Kassambara, 2017). *K-means clustering* merupakan metode yang paling sering digunakan dalam *unsupervised learning* untuk mempartisi *dataset* yang dianalisis menjadi k kelompok, dimana k merepresentasikan jumlah kelompok yang telah ditentukan sebelum dilakukan analisis *clustering* (Kassambara, 2017). Metode *K-means* memiliki beberapa kelebihan yaitu sangat simpel dan memiliki kemampuan untuk mengelompokkan grup dalam jumlah besar dalam waktu yang relatif cepat juga efisien (Khotimah dkk., 2016). Namun seiring berkembangnya jaman, analisis data menggunakan *clustering* semakin berkembang yaitu dengan memperhitungkan nilai derajat keanggotaan dari setiap objek, hal tersebut merupakan bagian dari himpunan *fuzzy* yang digunakan sebagai dasar dari pembobotan pada pengelompokan objek disebut *fuzzy clustering* (Bezdek, 1981). Pengelompokan pada *fuzzy clustering* dilakukan dengan mengelompokkan objek berdasarkan kedekatan dari nilai derajat keanggotaan yang dimiliki setiap objek yang mana nilai derajat keanggotaan tersebut terletak diantara rentang 0 dan 1 (Kakarla & Rama Mohan Babu, 2019). Salah satu metode *fuzzy clustering* yang paling sering digunakan adalah *Fuzzy C-means* (Dubey dkk., 2018). Keunggulan *Fuzzy C-means* (FCM) yaitu termasuk kedalam *soft clustering* sehingga sebuah objek yang berada pada batasan beberapa *cluster* dapat masuk kedalam beberap *cluster* pula tidak dipaksa harus masuk hanya kedalam satu *cluster* (Bora dkk., 2014). Pada kedua metode *clustering* tersebut yaitu *K-means* dan FCM menggunakan perhitungan jarak yang sama yaitu menggunakan perhitungan jarak Euclidean (Javed Mehedi Shamrat dkk., 2020; Lei

dkk., 2018). Jarak Euclidean merupakan jarak yang paling sering digunakan dalam metode *clustering*, dikarenakan dengan menggunakan jarak tersebut memiliki sebuah kelebihan yaitu jarak antar dua objek tidak akan berubah meskipun terdapat penambahan objek baru ke dalam analisis tersebut (Bora & Gupta, 2014).

Berdasarkan penjelasan diatas, artikel ini akan fokus membahas tentang pengelompokan provinsi di Indonesia menurut data intensitas bencana alam tahun 2017-2020 menggunakan *K-means* dan FCM. Terdapat beberapa penelitian yang menggunakan metode *clustering* dengan menggunakan metode *K-means* dan FCM telah dilakukan oleh (Mohammadrezapou dkk., 2020) untuk mengidentifikasi daerah *homogeneous* berdasarkan kualitas air tanahnya. Kemudian juga terdapat penelitian lainnya yang menggunakan *K-means* dan FCM oleh (Zeraatpisheh dkk., 2019) yaitu untuk memisahkan dan memperbarui peta tanah warisan di Iran Tengah. Hasil dari artikel ini diharapkan dapat menjadi referensi ataupun pedoman bagi pemerintah di Indonesia dalam membuat kebijakan di bidang penanggulangan bencana berdasarkan jumlah intensitas bencana alam yang terjadi di Indonesia.

KAJIAN TEORI

BENCANA ALAM

Bencana alam diartikan sebagai fenomena ekologis yang tiba-tiba terjadi dengan kekuatan yang cukup besar dimana penanggannya diperlukan bantuan dari pihak luar (Noji, 1991). Bencana alam juga dapat didefinisikan sebagai beberapa dampak yang terjadi pada lingkungan alam dengan skala besar dan cepat, dengan intensitas yang seketika juga mendalam terhadap sistem sosial-ekonomi (Alexander, 2017). Akibat yang terjadi dikarenakan oleh bencana alam dapat bervariasi, salah satunya yaitu dampaknya terhadap ekonomi suatu wilayah atau komunitas tersebut. Berdasarkan penelitian yang telah dilakukan oleh (Cavallo & Noy, 2011) dalam menganalisis dampak dari bencana alam terhadap bidang ekonomi, diketahui bahwa bencana alam dapat mengakibatkan beberapa kerusakan yaitu kerusakan perumahan, kerusakan tanaman, kerusakan infrastruktur, dan kerusakan manufaktur di sekitar daerah yang terdampak bencana alam.

Indonesia merupakan salah satu negara dengan intensitas bencana alam terbanyak di dunia setelah

Amerika Serikat (UCLouvain dkk., 2019). Hal tersebut dikarenakan letak geografis Indonesia yang berada pada gugus tektonik yaitu *ring of fire*, dimana hal tersebut membuat Indonesia memiliki intensitas bencana alam yang cukup tinggi (Pambudi, 2018). Pada tahun 2020, Indonesia tercatat memiliki jumlah intensitas bencana alam sebesar 4.977 kasus, dimana jumlah tersebut merupakan jumlah terbanyak kasus bencana alam yang pernah terjadi di Indonesia (BNPB, 2021a). Jumlah tersebut merupakan jumlah yang sangat besar, sehingga perlu dilakukan tindakan penanggulangan untuk menanggulangi dampak dari terjadinya bencana alam.

ELBOW METHOD

Elbow *method* adalah sebuah metode optimalisasi yang paling sering digunakan dalam mengidentifikasi jumlah *cluster* optimal dalam suatu *dataset* (Masud dkk., 2018). Pengertian lain dari Elbow *method* yaitu metode tertua untuk mengidentifikasi jumlah *cluster* optimal pada *dataset* yang dianalisis, yang mana ide dasar dari metode tersebut yaitu menentukan $k = 2$ sebagai jumlah *cluster* optimal awal k . Kemudian, jumlah *cluster* optimal k tersebut akan terus bertambah dari langkah pertama ke langkah yang paling maksimum untuk menentukan perkiraan jumlah *cluster* optimal potensial dan akhirnya akan dibedakan jumlah *cluster* optimal k sesuai dengan levelnya (Shi dkk., 2021).

K-MEANS

K-means merupakan sebuah metode yang digunakan pada *data mining* untuk melakukan proses pemodelan tanpa variabel kontrol dan juga salah satu metode yang digunakan untuk mengelompokkan data dalam partisi (Widiyaningtyas dkk., 2017). K-means juga merupakan algoritma *clustering* yang membagi setiap objek ke dalam beberapa *cluster* (Javed Mehedi Shamrat dkk., 2020).

Pada metode K-means membagi himpunan $X = \{X_1, X_2, \dots, X_d\}$ menjadi k *cluster* dengan $X_j = \{x_{1j}, x_{2j}, \dots, x_{tj}\}$ yang merepresentasikan objek ke- j untuk $j = \{1, 2, \dots, d\}$ dan *centroid* untuk cluster ke- i yaitu v_i dengan $i = \{1, 2, \dots, k\}$. Algoritma K-means didasarkan untuk meminimalkan jarak Euclidean dari suatu objek terhadap *centroid* yang didefinisikan sebagai berikut (L. M. Wang & Shao, 2018)

$$d_{ij} = \sqrt{\sum_{p=1}^d (v_{pi} - x_{pj})^2} \quad (1)$$

d_{ij} merupakan jarak Euclidean antara *centroid* ke- i dengan objek ke- j . Dengan v_{pi} merupakan *centroid* ke- i pada variabel ke- p , sedangkan untuk x_{pj} merupakan objek ke- j pada variabel ke- p .

$$v_i = \frac{\sum_{j=1}^d x_{tj}}{d} \quad (2)$$

Dengan v_i merupakan *centroid* ke- i dan x_j merupakan nilai dari objek ke- j . Kemudian dibagi dengan jumlah data pada setiap kelompok yaitu d .

FUZZY C-MEANS (FCM)

FCM adalah proses untuk menentukan nilai dari derajat keanggotaan setiap objeknya dan kemudian akan digunakan untuk menetapkan objek tersebut ke satu *cluster* atau lebih (Nithila & Kumar, 2016). Pada metode FCM membagi sebuah himpunan dari t vektor dimensi $X = \{X_1, X_2, \dots, X_d\}$ menjadi k *cluster* dengan $X_j = \{x_{1j}, x_{2j}, \dots, x_{tj}\}$ merepresentasikan sampel ke- j untuk $j = 1, 2, \dots, d$. Untuk sampel ke- j yaitu X_j dan *centroid cluster* ke- i yaitu v_i terdapat nilai derajat keanggotaan u_{ij} yang menunjukkan derajat keanggotaan dari sampel X_j terhadap *centroid* $v_i = v_{1i}, v_{2i}, \dots, v_{ti}$ untuk $i = 1, 2, \dots, k$. Dimana akan menghasilkan matriks partisi $U = (u_{ij})_{d \times k}$. Algoritma dari FCM didasarkan untuk meminimalisir fungsi objektif F_{fuzz} yang didefinisikan sebagai berikut.

$$F_{fuzz} = \sum_{j=1}^d \sum_{i=1}^k u_{ij}^m d_{ji}^2 \quad (3)$$

Dengan d_{ji} merupakan jarak Euclidean antara X_j terhadap *centroid cluster* v_i , yang dapat dihitung menggunakan rumus berikut.

$$d_{ij} = \sqrt{\sum_{p=1}^d (v_{pi} - x_{pj})^2} \quad (4)$$

Parameter m pada persamaan (3) merupakan parameter *fuzzifier* dan parameter tersebut mendefinisikan *fuzziness* pada *clustering*, dimana $m > 1$ atau biasanya diambil $m = 2$ (Babuška dkk., 2002). Rumus untuk menghitung derajat keanggotaan u_{ij} dan *centroid* v_i , dapat dituliskan sebagai berikut.

$$u_{ij} = \frac{1}{\sum_{p=1}^k \left(\frac{d_{ji}}{d_{jp}}\right)^{\frac{2}{m-1}}} \quad (5)$$

$$v_i = \frac{\sum_{j=1}^d u_{ij}^m X_j}{\sum_{j=1}^d u_{ij}^m} \quad (6)$$

Untuk pemilihan parameter $m \neq 1$ serta untuk nilai $i = 1, 2, \dots, k$.

SILHOUETTE COEFFICIENT

Dalam metode *clustering*, metode yang digunakan dalam melakukan validasi yaitu menggunakan *Silhouette Coefficient*. Metode *Silhouette* diperkenalkan pertama kali oleh (Rousseeuw, 1987) untuk menentukan kualitas dari hasil *clustering* yang telah dilakukannya. Metode *Silhouette* menunjukkan representasi garis dari setiap objek yang telah dikelompokkan dalam suatu *cluster* (Subbalakshmi dkk., 2015).

$$S_j = \frac{b_{qj} - a_{ij}}{\max\{a_{ij}, b_{qj}\}} \quad (7)$$

Dimana a_{ij} adalah jarak rata-rata dari objek j terhadap semua objek lainnya yang berada pada *cluster* i . Sedangkan, untuk b_{qj} jarak rata-rata dari objek j terhadap semua objek lainnya yang berada pada *cluster* q , dengan $i \neq q$ (Campello & Hruschka, 2006).

Hasil dari *Silhouette Coefficient* berkisar diantara rentang -1 sampai dengan $+1$. Jika Hasil dari *Silhouette Coefficient* adalah $+1$, maka menunjukkan objek tersebut dikelompokkan dengan benar pada *cluster* tersebut. Namun, jika nilai dari *Silhouette Coefficient* berada didekat -1 maka objek ke- j tersebut berada pada *cluster* yang salah, sehingga objek ke- j akan lebih baik jika di kelompokkan pada *cluster* yang berdekatan dengan *cluster* tersebut (Subbalakshmi dkk., 2015).

Dalam metode *Silhouette* terdapat berapa kategori subjektif dalam melakukan pengukuran pengelompokan menurut (Kaufman & Rousseeuw, 1990), dapat dilihat pada tabel sebagai berikut.

Tabel 1 Kategori Silhouette Coefficient

Silhouette Coefficient	Kategori
0.71 - 1.00	Kategori kuat
0.51 - 0.70	Kategori baik
0.26 - 0.50	Kategori lemah
≤ 0.25	Kategori buruk

METODE

DATA PENELITIAN

Data yang digunakan dalam artikel ini adalah data sekunder dari Data Informasi Bencana Indonesia yang dikelola oleh Badan Nasional Penanggulangan bencana (BNPB) yang berupa data intensitas bencana alam yang terjadi di Indonesia pada tahun 2017 - 2021, data tersebut dapat diakses melalui *website* resmi BNPB. Pada data intensitas bencana tersebut terdiri dari 8 jenis bencana alam yang terjadi di 34 provinsi yang ada di Indonesia. Bencana alam tersebut meliputi banjir, tanah longsor, abrasi, angin puting beliung, kekeringan, kebakaran hutan dan lahan, gempa bumi, dan letusan gunung api.

ALGORITMA K-MEANS

K-means merupakan pengolahan data dengan algoritma yang cukup mudah dimana mempartisi *dataset* menjadi beberapa *cluster* k . Algoritma K-means tersebut cukup mudah untuk diimplementasikan, relatif cepat, dan sangat sering digunakan (Silitonga, 2017).

Dalam pengolahan data menggunakan metode *clustering* K-means menurut (Suwanda dkk., 2020), perlu dilakukan beberapa langkah antara lain yaitu, langkah pertama, menentukan jumlah *cluster* k yang diinginkan. Kemudian, langkah kedua yaitu menentukan *centroid* untuk setiap *cluster*. Pada langkah kedua ini *centroid* ditentukan dengan cara memilih data secara acak sejumlah *cluster* yang diinginkan. Namun, pada saat iterasi cara menentukan *centroid* untuk setiap *cluster* yaitu dengan menggunakan persamaan (2). Langkah ketiga yaitu menentukan jarak Euclidean untuk setiap objek terhadap setiap *cluster* menggunakan persamaan (1).

Langkah keempat yaitu akan dikelompokkan objek tersebut kedalam suatu *cluster* dengan cara meninjau jarak terpendek objek tersebut terhadap suatu *centroid*. Langkah kelima yaitu melakukan pengecekan keadaan berhenti yaitu pada saat tidak ada objek yang berpindah *cluster*. Namun, bila terjadi sebaliknya maka algoritma tersebut diulang kembali dari langkah kedua yaitu menentukan *centroid* menggunakan persamaan (2) dan akan terus berulang hingga mencapai keadaan berhenti atau konvergen.

ALGORITMA FUZZY C-MEANS (FCM)

FCM merupakan bentuk modifikasi dari metode K-means yang mana bekerja berdasarkan logika *fuzzy* untuk menentukan fungsi keanggotaan dan nilai derajat keanggotaan dari setiap objek (Mohammadrezapou dkk., 2020).

Dalam pengolahan data menggunakan FCM *clustering* menurut (Tao dkk., 2018) perlu dilakukan beberapa langkah antara lain yaitu, langkah pertama, menentukan jumlah *cluster* k yang diinginkan, derajat *fuzziness* m , dan nilai *error* yang diharapkan ϵ . Kemudian langkah kedua yaitu menentukan nilai derajat keanggotaan dari setiap objek menggunakan persamaan (8) sebagai berikut (Z. Wang dkk., 2012).

$$\bar{x}_{tj} = \frac{x_{tj}}{x(\text{maks})_j} \tag{8}$$

Langkah ketiga yaitu menentukan matriks partisi U secara acak dengan ordo $n \times j$, dengan syarat jumlah elemen setiap barisnya harus sama dengan 1. Langkah keempat yaitu *centroid* v_i dengan $i = 1, 2, \dots, k$ untuk setiap *cluster* dengan menggunakan persamaan (6). Setelah diketahui *centroid* untuk setiap *cluster* v_i , langkah kelima yaitu menentukan jarak Euclidean untuk setiap objek X_j terhadap *centroid* v_i dengan menggunakan persamaan (4).

Langkah keenam yaitu menentukan semua nilai derajat keanggotaan u_{ij} dengan menggunakan persamaan (5). Langkah ketujuh yaitu memperbarui matriks partisi U yang telah ada sebelumnya dengan nilai setiap elemennya yaitu nilai derajat keanggotaan u_{ij} yang telah dihitung sebelumnya dengan persamaan (5). Langkah kedelapan yaitu menentukan fuungsi objektif F_{fuzz} dengan menggunakan persamaan (3).

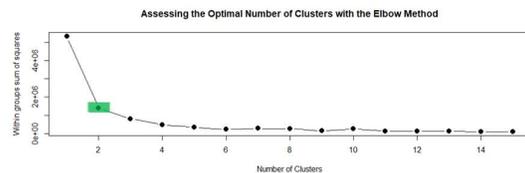
Langkah kesembilan yaitu melakukan pengecekan untuk keadaan berhenti yaitu jika $|F_{fuzz(n)} - F_{fuzz(n-1)}| < \epsilon$, maka algoritma FCM tersebut akan berhenti dan telah mencapai keadaan konvergen. Namun, jika sebaliknya maka algoritma FCM tersebut akan diulang kembali dari langkah kedua yaitu pada langkah menentukan matriks partisi U dan akan terus berulang hingga mencapai keadaan berhenti atau konvergen.

HASIL DAN PEMBAHASAN

Dalam menganalisis suatu data, perlu diketahui terlebih dahulu kondisi awal dari data yang akan dianalisis. Kondisi awal dari suatu data dapat diketahui dengan menggunakan analisis statistik,

sehingga dapat menunjukkan analisis deskriptif dari data tersebut. Tujuan dari proses tersebut adalah untuk mengetahui karakteristik dari setiap variabel intensitas bencana alam yang terjadi di Indonesia setiap tahunnya dengan meninjau nilai *means*, nilai minimum dan nilai maksimum. Berdasarkan hasil analisis statistik diperoleh bahwa nilai *means* tertinggi dimiliki oleh bencana banjir tahun 2020 dengan nilai sebesar 44.68, dengan nilai maksimum sebesar 254 kasus yang terjadi di Jawa Tengah dan nilai minimum sebesar 2 kasus yang terjadi di Kepulauan Riau. Sedangkan, untuk nilai *means* terendah dimiliki oleh bencana letusan gunung api tahun 2021 dengan nilai sebesar 0.08824, dengan nilai maksimum sebesar 2 kasus yang terjadi di Sumatra Utara dan nilai minimum sebesar 0 kasus yaitu terjadi hampir di seluruh Indonesia, kecuali Provinsi Sumatra Utara dan Jawa Timur. Jenis bencana alam yang termasuk dalam kategori tinggi (*means* berada di rentang 32.76-44.68) yaitu bencana banjir tahun 2021, bencana angin puting beliung 2018, bencana tanah longsor tahun 2020, bencana angin puting beliung tahun 2019, bencana angin puting beliung tahun 2020, dan bencana banjir tahun 2020. Sedangkan, jenis bencana alam lainnya termasuk dalam kategori sedang ataupun rendah.

Langkah pertama yang perlu dilakukan pada metode *clustering* yaitu menentukan jumlah *cluster* optimal k yang akan digunakan dalam metode *clustering* selanjutnya. Pada artikel ini akan digunakan *Elbow method* untuk menentukan jumlah *cluster* optimalnya. Dengan melakukan pengolahan data menggunakan bantuan *software* Rstudio diketahui bahwa jumlah *cluster* optimal untuk data intensitas bencana alam di Indonesia tahun 2017-2021 yaitu 2 *cluster*. Hasil pengolahan data dengan *Elbow method* dapat dilihat seperti pada Gambar 1 berikut.



Gambar 1 Jumlah Cluster Optimal Menggunakan Elbow Method

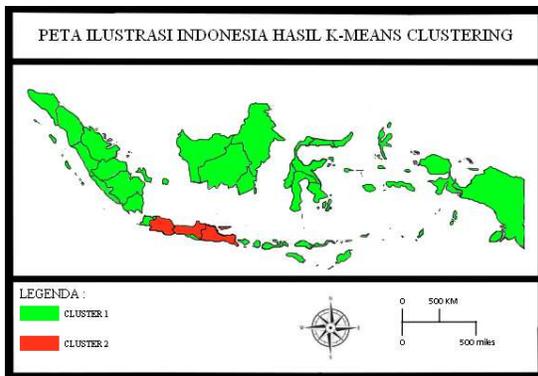
Langkah selanjutnya dalam metode *clustering* setelah diketahui jumlah *cluster* optimal adalah melakukan proses *clustering* menggunakan metode

K-means dan FCM pada data intensitas bencana alam di Indonesia tahun 2017-2021.

Metode pertama yang digunakan adalah metode K-means *clustering*. K-means merupakan metode *clustering* yang diperoleh melalui perhitungan jarak setiap objek pada data dengan *centroid* pada setiap *cluster*. Perhitungan *centroid* untuk setiap *cluster* pada K-means dapat dihitung menggunakan persamaan (2). Setelah diketahui *centroid* untuk setiap *cluster*, kemudian akan dihitung jarak untuk setiap objek pada data terhadap *centroid* pada setiap *cluster* dengan menggunakan persamaan (1).

Langkah selanjutnya yaitu objek-objek tersebut akan dikelompokkan objek tersebut kedalam suatu *cluster* dengan cara memperhatikan jarak terpendek objek tersebut terhadap suatu *centroid*. Setelah itu, akan dilakukan pengecekan kondisi berhenti atau konvergen yaitu pada saat tidak ada objek yang berpindah *cluster*. Namun, jika terjadi keadaan sebaliknya maka algoritma K-means *clustering* akan diulang kembali hingga mencapai keadaan berhenti atau konvergen.

Berdasarkan hasil optimasi menggunakan Elbow *method* diketahui bahwa jumlah *cluster* optimal dalam pengelompokan menggunakan K-means *clustering* adalah 2 *cluster*. Maka hasil pengelompokan menggunakan algoritma K-means *clustering* dengan jumlah *cluster* optimal yaitu 2 *cluster*, dapat direpresentasikan dalam Gambar 2 sebagai berikut.



Gambar 2 Peta Ilustrasi Indonesia Hasil K-means Clustering

Dari Gambar 2 dapat diketahui bahwa pengelompokan menggunakan metode K-means *clustering* membagi provinsi di Indonesia menjadi 2 *cluster*. Dengan *cluster* 1 yang dilambangkan dengan warna hijau, yang beranggotakan 31 provinsi yaitu Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi,

Sumatra Selatan, Bengkulu, Lampung, Kep Bangka Belitung, Kepulauan Riau, DKI Jakarta, Di Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua. Sedangkan, untuk *cluster* 2 dilambangkan dengan warna merah, yang beranggotakan 3 Provinsi yaitu Jawa Barat, Jawa Tengah, dan Jawa Timur.

Metode kedua yaitu menggunakan FCM *clustering*. FCM *clustering* merupakan salah satu metode dalam *fuzzy clustering* yang mana prosesnya sedikit berbeda dengan *clustering* pada umumnya. Perbedaan pada metode FCM ialah perlu menentukan nilai derajat keanggotaan untuk setiap objek pada data terlebih dahulu sebelum melakukan proses *clustering*. Nilai derajat keanggotaan untuk setiap objek dapat dihitung menggunakan persamaan (8). Kemudian akan ditentukan matriks partisi U untuk data tersebut yang berordo $n \times j$ secara acak dengan syarat jumlah elemen setiap barisnya harus sama dengan 1.

Langkah selanjutnya yaitu menentukan *centroid* v_i untuk setiap *cluster* dengan menggunakan persamaan (6). Kemudian, akan dihitung jarak Euclidean dari setiap objek terhadap *centroid* setiap *cluster* dengan menggunakan persamaan (4). Setelah itu, akan ditentukan derajat keanggotaan dari setiap objek terhadap *centroid* pada setiap *cluster* menggunakan persamaan (5). Langkah selanjutnya yang akan dilakukan yaitu menentukan fungsi objektif dari data tersebut menggunakan persamaan (3). Kemudian, akan dilakukan pengecekan kondisi berhenti yaitu jika memenuhi $|F_{fuzz(n)} - F_{fuzz(n-1)}| < \epsilon$, maka algoritma FCM tersebut akan berhenti dan telah mencapai keadaan konvergen. Namun, jika sebaliknya maka algoritma FCM tersebut akan diulang hingga mencapai keadaan berhenti atau konvergen.

Berdasarkan hasil optimasi menggunakan Elbow *method* diketahui bahwa jumlah *cluster* optimal dalam pengelompokan menggunakan FCM *clustering* adalah 2 *cluster*. Maka hasil pengelompokan menggunakan algoritma FCM *clustering* dengan jumlah *cluster* optimal yaitu 2 *cluster*, dapat direpresentasikan dalam Gambar 3 sebagai berikut.



Gambar 3 Peta Ilustrasi Indonesia Hasil FCM Clustering

Dari Gambar 3 dapat diketahui bahwa pengelompokan menggunakan metode FCM clustering membagi provinsi di Indonesia menjadi 2 cluster. Dengan cluster 1 yang dilambangkan dengan warna hijau, yang beranggotakan 32 provinsi di Indonesia yaitu Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kep Bangka Belitung, Kepulauan Riau, DKI Jakarta, Di Yogyakarta, Banten, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua. Sedangkan, untuk cluster 2 dilambangkan dengan warna merah, yang beranggotakan 2 provinsi yaitu Jawa Barat dan Jawa Tengah.

Setelah dilakukan proses clustering menggunakan K-means dan FCM, hasil dari proses clustering tersebut perlu dilakukan validasi untuk mengetahui kualitas dan akurasi dari hasil tersebut. Pada artikel ini akan fokus menggunakan *Silhouette Coefficient* dalam melakukan validasi. Hasil validasi menggunakan *Silhouette Coefficient* untuk metode clustering K-means dan FCM, dapat ditinjau pada tabel berikut.

Tabel 2 Hasil Rata-Rata *Silhouette Coefficient*

Metode	Rata-rata <i>Silhouette Coefficient</i>	Kategori
K-means	0.8031255	Kategori kuat
FCM	0.8403194	Kategori kuat

Dari hasil *Silhouette Coefficient* pada Tabel 2, dapat diketahui bahwa dengan menggunakan metode K-

means clustering memiliki rata-rata *Silhouette Coefficient* sebesar 0.8031255 yang termasuk dalam kategori pengelompokan yang kuat, sedangkan dengan menggunakan FCM clustering memiliki rata-rata *Silhouette Coefficient* sebesar 0.8403194 yang termasuk dalam kategori pengelompokan yang kuat. Dari kedua metode clustering tersebut dapat diketahui bahwa dengan menggunakan FCM clustering memiliki rata-rata *Silhouette Coefficient* yang lebih besar dibandingkan menggunakan K-means clustering. Sehingga dapat dikatakan bahwa dengan menggunakan metode FCM clustering hasil pengelompokan yang dihasilkan memiliki akurasi yang lebih baik dibandingkan dengan menggunakan K-means clustering. Maka dapat disimpulkan bahwa pengelompokan dengan nilai akurasi terbaik dalam mengelompokkan data intensitas bencana alam di Indonesia tahun 2017-2021 yaitu menggunakan metode FCM clustering.

Hasil dari artikel ini merupakan pengembangan metode yang telah dilakukan dari penelitian sebelumnya yang dilakukan oleh (Yana dkk., 2018), dimana pada penelitian tersebut menggunakan metode K-means untuk melakukan pengelompokan datanya. Sehingga pada artikel ini dilakukan pengembangan metode yaitu dengan menggunakan *fuzzy clustering* berupa FCM dan didapatkan hasil bahwa dengan menggunakan FCM hasil pengelompokan yang dihasilkan lebih baik dibandingkan dengan menggunakan K-means yang mana dapat dilihat dari rata-rata *Silhouette Coefficient* FCM lebih besar dibandingkan dengan K-means yaitu 0.8403194 yang termasuk dalam kategori pengelompokan yang kuat.

Dengan meninjau hasil *Silhouette Coefficient* diketahui bahwa FCM clustering merupakan metode yang memiliki akurasi paling baik dalam mengelompokkan data intensitas bencana alam di Indonesia tahun 2017-2021. Maka dari berdasarkan hasil clustering dengan menggunakan FCM clustering provinsi di Indonesia dikelompokkan menjadi 2 cluster. Dengan cluster 1 dapat dikatakan dengan provinsi yang tidak rawan terjadi bencana alam. Hal tersebut dikarenakan pada provinsi yang termasuk cluster 1 memiliki intensitas bencana alam yang relatif rendah, sehingga dapat dikatakan bahwa sebagai provinsi yang tidak rawan terjadi bencana alam. Sedangkan untuk cluster 2 dapat dikatakan dengan provinsi yang rawan terjadi bencana alam. Hal

tersebut dikarenakan pada provinsi yang termasuk dalam *cluster* 2 memiliki intensitas bencana alam yang sangat tinggi dibandingkan dengan *cluster* 1, sehingga dapat dikatakan sebagai provinsi yang rawan terjadi bencana alam.

Analisis *cluster* menggunakan metode K-means *clustering* dan FCM *clustering* bertujuan untuk mengelompokkan provinsi di Indonesia yang mana diharapkan dapat menjadi referensi ataupun pedoman bagi pemerintah di Indonesia dalam membuat kebijakan berdasarkan jumlah intensitas bencana alam yang terjadi di Indonesia.

PENUTUP

SIMPULAN

Pada artikel ini, metode pengelompokan K-means dan FCM bertujuan untuk mengelompokkan negara-negara di Asia berdasarkan data jumlah kasus baru kanker pada tahun 2020, yang mana hal tersebut diharapkan dapat menjadi referensi atau pedoman bagi pemerintah di Asia untuk membuat kebijakan di bidang penanggulangan bencana berdasarkan jumlah intensitas bencana alam yang terjadi di Indonesia.

Berdasarkan proses analisis *cluster* optimal menggunakan Elbow *method* diketahui bahwa jumlah *cluster* optimal dari data intensitas bencana alam di Indonesia tahun 2017-2021 yaitu 2 *cluster*. Kemudian setelah dilakukan analisis menggunakan metode K-means dan FCM menggunakan *Silhouette Coefficient* diketahui bahwa rata-rata *Silhouette Coefficient* metode FCM memiliki nilai akurasi yang lebih baik dibandingkan dengan K-means yaitu sebesar 0.8403194 yang termasuk dalam kategori kuat.

Maka hasil *clustering* dengan menggunakan metode FCM *clustering* yaitu data intensitas bencana alam di Indonesia tahun 2017-2021 terbagi menjadi 2 *cluster*. Pada *cluster* 1 yang dapat dikatakan sebagai provinsi yang tidak rawan terjadi bencana alam memiliki 32 anggota yaitu Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kep Bangka Belitung, Kepulauan Riau, DKI Jakarta, Di Yogyakarta, Banten, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo,

Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua.

Sedangkan, pada *cluster* 2 yang dapat dikatakan sebagai provinsi yang rawan terjadi bencana alam memiliki 2 anggota yaitu Jawa Barat dan Jawa Timur.

SARAN

Untuk penelitian lebih lanjut dapat dilakukan menggunakan metode *clustering* lainnya berupa metode *hierarchical*, metode *non-hierarchical*, dan metode *fuzzy clustering* lainnya. Salah satu contohnya yaitu metode *Fuzzy Gustafson-Kessel*, *Fuzzy K-medoid*, dan *K-medoids*.

DAFTAR PUSTAKA

- Alexander, D. (2017). Natural Disaster. In *Angewandte Chemie International Edition*, 6(11), 951–952. Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9781315859149/natural-disasters-david-alexander>
- Babuška, R., Van Der Veen, P. J., & Kaymak, U. (2002). Improved covariance estimation for Gustafson-Kessel clustering. *IEEE International Conference on Fuzzy Systems*, 2(usually 1), 1081–1085. <https://doi.org/10.1109/fuzz.2002.1006654>
- Bateni, M. H., Behnezhad, S., Derakhshan, M., Hajiaghayi, M. T., Kiveris, R., Lattanzi, S., & Mirrokni, V. (2017). Affinity clustering: Hierarchical clustering at scale. *Advances in Neural Information Processing Systems*, 2017-Decem(Nips), 6865–6875.
- Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. A Division of Plenum Publishing Corporation.
- BNPB, B. N. P. B. (2021a). *Data Informasi Bencana Alam*. Pusat Data Informasi dan Komunikasi Kebencanaan (Pusdatinkom). <https://dibi.bnpb.go.id/>
- BNPB, B. N. P. B. (2021b). *Data Informasi Bencana Indonesia Erupsi Gunung Semeru*. Badan Nasional Penanggulangan Bencana. <https://dibi.bnpb.go.id/xdibi/read/54875/35//111/2021//2//1>
- Bora, Jyoti, D., Gupta, & Kumar, D. A. (2014). A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm. *International Journal of Computer Trends and Technology*, 10(2), 108–113. <https://doi.org/10.14445/22312803/ijctt-v10p119>
- Bora, M. D. J., & Gupta, D. A. K. (2014). *Effect of Different Distance Measures on the Performance of K-Means Algorithm: An Experimental Study in*

- Matlab*. 5(2), 2501–2506.
<http://arxiv.org/abs/1405.7471>
- Campello, R. J. G. B., & Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems*, 157(21), 2858–2875.
<https://doi.org/10.1016/j.fss.2006.07.006>
- Cavallo, E., & Noy, I. (2011). Natural disasters and the economy - A survey. *International Review of Environmental and Resource Economics*, 5(1), 63–102. <https://doi.org/10.1561/101.00000039>
- Chaudhary, M. T., & Piracha, A. (2021). *Natural Disasters – Origins , Impacts , Management*. 1101–1131.
- Dubey, A. K., Gupta, U., & Jain, S. (2018). Comparative study of K-means and fuzzy C-means algorithms on the breast cancer data. *International Journal on Advanced Science, Engineering and Information Technology*, 8(1), 18–29. <https://doi.org/10.18517/ijaseit.8.1.3490>
- Glossary, H. (2014). IRDR DATA Reports - Perils Classification and Hazards Glossary. *Disaster Risk IPO*.
- Hall, R. (2009). Geology of Indonesia. *Geology of Indonesia*, 454–460.
- Hillis, R. R., & Müller, R. . (2003). *Evolution and Dynamics of the Australian Plate*. Geological Society of America. https://books.google.co.id/books?id=Ch_1JILkKE0C&q=new+guinea+north+west+tectonics&pg=PA4&redir_esc=y#v=snippet&q=new+guinea+north+west+tectonics&f=false
- IFRC. (2021). *What is disaster?* IFRC. <https://www.ifrc.org/what-disaster>
- Javed Mehedi Shamrat, F. M., Tasnim, Z., Mahmud, I., Jahan, N., & Nobel, N. I. (2020). Application of k-means clustering algorithm to determine the density of demand of different kinds of jobs. *International Journal of Scientific and Technology Research*, 9(2), 2550–2557.
- Johnson, D. (2004). *The Geology of Australia* (C. U. PRESS (ed.)). THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE.
- Kakarla, N. M. L., & Rama Mohan Babu, G. (2019). Application of fuzzy K-means (FKM) algorithms in identifying better clusters of few drugs from drugbank database. *International Journal of Innovative Technology and Exploring Engineering*, 8(6 Special Issue 4), 655–660. <https://doi.org/10.35940/ijitee.F1134.0486S419>
- Kassambara, A. (2017). *Practical Guide To Cluster Analysis in R*. http://www.sthda.com/sthda/ebooks/clustering_english_edition1_preview.pdf
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Group in Data: An Introduction to Cluster Analysis*. A John Wiley & Sons, Inc., Publication.
- Khotimah, B. K., Irhamni, F., & Sundarwati, T. (2016). A genetic algorithm for optimized initial centers K-means clustering in SMEs. *Journal of Theoretical and Applied Information Technology*, 90(1), 23–30.
- Lei, T., Jia, X., Zhang, Y., He, L., Meng, H., & Nandi, A. K. (2018). Significantly Fast and Robust Fuzzy C-Means Clustering Algorithm Based on Morphological Reconstruction and Membership Filtering. *IEEE Transactions on Fuzzy Systems*, 26(5), 3027–3041. <https://doi.org/10.1109/TFUZZ.2018.2796074>
- Masud, M. A., Huang, J. Z., Wei, C., Wang, J., Khan, I., & Zhong, M. (2018). I-nice: A new approach for identifying the number of clusters and initial cluster centres. *Information Sciences*, 466, 129–151. <https://doi.org/10.1016/j.ins.2018.07.034>
- Masyekh, & Faradilla, A. (2021). Signifikansi Pendataan Pengungsi Erupsi Semeru di Lumajang. *Badan Nasional Penanggulangan Bencana*. <https://bnpb.go.id/berita/signifikansi-pendataan-pengungsi-erupsi-semeru-di-lumajang>
- Mohammadrezapou, Kisi, O., Pourahmad, O., & Fariba. (2020). Fuzzy c-means and K-means clustering with genetic algorithm for identification of homogeneous regions of groundwater quality. *Neural Computing and Applications*, 32(8), 3763–3775. <https://doi.org/10.1007/s00521-018-3768-7>
- Nithila, E. E., & Kumar, S. S. (2016). Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering. *Alexandria Engineering Journal*, 55(3), 2583–2588. <https://doi.org/10.1016/j.aej.2016.06.002>
- Noji, E. K. (1991). Natural disasters. *Critical Care Clinics*, 7(2), 271–292. [https://doi.org/10.1016/s0749-0704\(18\)30306-3](https://doi.org/10.1016/s0749-0704(18)30306-3)
- Omran, M. G. H., Engelbrecht, A. P., & Salman, A. (2007). An overview of clustering methods. *Intelligent Data Analysis*, 11(6), 583–605. <https://doi.org/10.3233/ida-2007-11602>
- Pambudi, N. A. (2018). Geothermal power generation in Indonesia, a country within the ring of fire: Current status, future development and policy. *Renewable and Sustainable Energy Reviews*, 81(March), 2893–2901. <https://doi.org/10.1016/j.rser.2017.06.096>
- Rousseeuw, P. J. (1987). *Silhouettes : a graphical aid to the interpretation and validation of cluster analysis*. 20, 53–65.

- Shi, C., Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *Eurasip Journal on Wireless Communications and Networking*, 2021(1). <https://doi.org/10.1186/s13638-021-01910-w>
- Silitonga, P. D. . (2017). Clustering of Patient Disease Data by Using K-Means Clustering. *International Journal of Computer Science and Information Security IJCSIS*, 15(7), 219-233.
- Subbalakshmi, C., Rama Krishna, G., Krishna Mohan Rao, S., & Venketeswa Rao, P. (2015). A method to find optimum number of clusters based on fuzzy silhouette on dynamic data set. *Procedia Computer Science*, 46(Icict 2014), 346-353. <https://doi.org/10.1016/j.procs.2015.02.030>
- Suwanda, R., Syahputra, Z., & Zamzami, E. M. (2020). Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K. *Journal of Physics: Conference Series*, 1566(1). <https://doi.org/10.1088/1742-6596/1566/1/012058>
- Tao, Y., Zhang, Y., & Wang, Q. (2018). Fuzzy c-mean clustering-based decomposition with GA optimizer for FSM synthesis targeting to low power. *Engineering Applications of Artificial Intelligence*, 68(October 2017), 40-52. <https://doi.org/10.1016/j.engappai.2017.10.022>
- UCLouvain, Disaster, C. for R. on the E. of, & USAID. (2019). *Natural Disaster*. https://cred.be/sites/default/files/adsr_2019.pdf
- Verstappen, H. T. (2010). Indonesian Landforms and Plate Tectonics. *Indonesian Journal on Geoscience*, 5(3), 197-207. <https://doi.org/10.17014/ijog.v5i3.103>
- Wang, L. M., & Shao, Y. M. (2018). Crack Fault Classification for Planetary Gearbox Based on Feature Selection Technique and K-means Clustering Method. *Chinese Journal of Mechanical Engineering (English Edition)*, 31(1). <https://doi.org/10.1186/s10033-018-0202-0>
- Wang, Z., Cui, T., Wang, Y., & Yu, Z. (2012). Flood season division with an improved fuzzy C-mean clustering method in the Taihu lake basin in China. *Procedia Engineering*, 28(2011), 66-74. <https://doi.org/10.1016/j.proeng.2012.01.684>
- WHO/EHA. (2002). Disasters & Emergencies; Definitions Training Package. *WHO/EHA Training package, March*, 1-26. <http://apps.who.int/disasters/repo/7656.pdf>
- Widiyaningtyas, T., Prabowo, M. I. W., & Pratama, M. A. M. (2017). Implementation of k-means clustering method to distribution of high school teachers. *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2017-Decem(September)*, 19-21. <https://doi.org/10.1109/EECSI.2017.8239083>
- Yana, M. S., Setiawan, L., Ulfa, E. M., & Rusyana, A. (2018). Penerapan Metode K-Means dalam Pengelompokan Wilayah Menurut Intensitas Kejadian Bencana Alam di Indonesia Tahun 2013-2018. *Journal of Data Analysis*, 1(2), 93-102. <https://doi.org/10.24815/jda.v1i2.12584>
- Yanuarto, T. (2021). Warga Mengungsi Akibat Letusan Gunung Semeru Sebanyak 3.697 Jiwa. *Badan Nasional Penanggulangan Bencana*. <https://bnpb.go.id/berita/-update-warga-mengungsi-akibat-letusan-gunung-semeru-sebanyak-3-697-jiwa>
- Zeraatpisheh, M., Ayoubi, S., Brungard, C. W., & Finke, P. (2019). Disaggregating and updating a legacy soil map using DSMART, fuzzy c-means and k-means clustering algorithms in Central Iran. *Geoderma*, 340(September 2018), 249-258. <https://doi.org/10.1016/j.geoderma.2019.01.005>