



*Corresponding author: Wahidah Sanusi, Department of Mathematics, Universitas Negeri Makassar, Makassar, 90223, Indonesia

E-mail: wahidah.sanusi@unm.ac.id

RESEARCH ARTICLE

The Implementation of Spatial Model with K-Means Clustering Method to Cluster Flood Affected Areas in Bone Regency

Irwan, Wahidah Sanusi*, Adil Saputra Anwar, & Abdul Rahman

Department of Mathematics, Universitas Negeri Makassar, Makassar, 90223, Indonesia

Abstract: This research is an applied research that aims to determine the clusters of areas affected by floods in bone regency. This study user the K-means method in clustering flood data in 2020 and 2021. The data is grouped based on the number of families affected and the resulting damage. In determining the number of clusters to be used, this study validated 3, 4 and 5 clusters using the davis bouldin index (DBI) validation in determining the best cluster. The results of this validation resulted in the best number of clusters, namely $k=4$ for 2020 data with a minimum DBI of 0,73 and $k=4$ for 2021 data with a minimum DBI of 0,44. After clustering, the cluster number for data for 2020 from the firs to fourth cluster member are 34, 7, 2 and 3 sequentially, while the data for years 2021 sequentially are 3, 22, 3 and 1. Then the cluster results are displayed in a spatial form that is created using ArcGIS.

Keywords: K-Means, cluster, flood, Bone Regency

1. Introduction

Clustering adalah suatu metode pengelompokan berdasarkan ukuran kedekatan (kemiripan). *Clustering* berbeda dengan group, kalau group berarti kelompok yang memiliki kondisi yang sama. Sedangkan *cluster* kelompoknya tidak harus sama, pengelompokan didasarkan pada kedekatan dari suatu karakteristik sampel yang ada, salah satunya dengan menggunakan rumus jarak *Euclidean* (Sanusi, dkk., 2020). Sebelum melakukan pengelompokan, maka dilakukan validasi dalam menentukan jumlah *cluster* yang akan digunakan. Salah satu cara dalam menentukan jumlah *cluster* yang akan digunakan adalah dengan validasi *Davies-Bouldin Index* (DBI). Sampai saat ini, para ilmuwan masih terus melakukan berbagai usaha untuk melakukan perbaikan model *cluster* dan menghitung jumlah *cluster* yang optimal sehingga dapat dihasilkan *cluster* yang paling baik. Ada dua metode yang kita kenal, yaitu *hierarchical Clustering* dan *partitioning*. Metode *hierarchical Clustering* sendiri terdiri dari *complete linkage Clustering*, *single linkage Clustering*, *average linkage Clustering* dan *centroid linkage Clustering*. Sedangkan metode *partitioning* sendiri terdiri dari *K-means* dan *fuzzy K-means* (Alfina & Santosa, 2012).

K-Means adalah salah *salu* metode dalam data mining dengan teknik unsupervised yang mengelompokkan data dengan sistem partitional (Syarif dkk., 2018). *K-Means* bekerja dengan mengelompokkan data yang ada ke dalam beberapa *cluster*. Hasil dari pengelompokan tersebut akan tampilkan dalam bentuk *spasial*. Analisis *spasial* merupakan analisis keruangan menggunakan data yang diinterpretasikan dalam bentuk peta. Analisis ini menggunakan alat bantu software terutama ArcGIS (Ali dan Trisutomo, 2017). Anasiru



(2016) menyatakan bahwa analisis *spasial* merupakan teknik atau proses yang terdiri atas sejumlah perhitungan dan evaluasi logika (matematis) dalam rangka menemukan potensi hubungan atau pola-pola. *Spasial* dapat diartikan sebagai sesuatu yang berkaitan dengan ruang atau tempat yang mengidentifikasi lokasi geografis kenampakan dan batas di bumi, salah satunya di Indonesia.

Indonesia merupakan daerah yang beriklim tropis dengan intensitas curah hujan cukup tinggi, sehingga hampir seluruh wilayah di Indonesia berpotensi terkena bencana banjir (Faiza, 2019). Fauziah (2008) menyatakan bahwa banjir merupakan genangan di suatu wilayah baik berupa dataran rendah maupun dataran tinggi yang disebabkan oleh curah hujan yang airnya tidak tertampung, terjadinya banjir disebabkan karena hujan merata di seluruh daerah dengan curah hujan yang tinggi. bahwa kepala Badan Penanggulangan Bencana Daerah (BPBD) Sulawesi Selatan menyatakan bahwa banjir antara lain terjadi di Kecamatan Biringkanaya, Tamalanrea, Manggala, Panakukang, Tamalate, dan Rappocini di Kota Makassar (Aditya, 2021). Salah satu kota di Provinsi Sulawesi Selatan yang rentan terjadi banjir adalah Kabupaten Bone.

Penelitian ini bertujuan untuk mengetahui daerah yang terdampak banjir di Kab. Bone dengan menampilkan gambaran *spasial*-nya. Salah satu metode yang digunakan untuk memperoleh gambaran daerah terdampak banjir di Kab. Bone adalah analisis *spasial* dengan metode *K-Means Clustering*.

2. Research Method and Materials

Jenis penelitian yang digunakan adalah penelitian terapan yang ditujukan untuk mendapatkan solusi dari suatu masalah yang ada di masyarakat, industri, pemerintahan sebagai kelanjutan dari riset dasar (Irina, 2017). Dalam penelitian ini, peneliti melakukan pengelompokan daerah terdampak banjir di Kabupaten Bone menggunakan metode *K-Means Clustering* dengan model spasial dalam menggambarkan titik wilayah banjir. Data yang digunakan adalah data sekunder yang diambil dari Kantor Badan Penanggulangan Bencana Daerah (BPBD) dan Badan Pusat Statistik (BPS) Kabupaten Bone Sulawesi Selatan. Pada penelitian ini data yang digunakan adalah data banjir yang terjadi pada tahun 2020 dan 2021.

2.1. Prosedur Pelaksanaan Penelitian

Prosedur pelaksanaan penelitian yang dilakukan adalah sebagai berikut:

- Mengumpulkan literatur tentang *K-Means Clustering* dan Validasi DBI
- Melakukan pengumpulan data sekunder bencana banjir di kantor BPBD Kabupaten Bone tahun 2020 dan 2021.
- Menentukan Variabel X yang digunakan yaitu X1 (jumlah kartu keluarga yang terdampak banjir), X2 (banyaknya unit fasilitas yang rusak), X3 (luas sawah keseluruhan yang rusak), X4 (panjang jalan yang rusak) dan X5 (luas kebun dan kolam yang terendam).
- Menentukan jumlah *Cluster* yang digunakan yaitu 3,4 dan 5
- Penerapan Metode *K-Means Clustering* (Fatmawati, 2018).
- Melakukan validasi *cluster* dengan menentukan nilai DBI dari *centroid* iterasi ke-n.
- Membandingkan DBI dari setiap *Cluster* untuk mendapatkan jumlah k optimal yang akan digunakan pada data tahun 2020 dan 2021.
- Membuat kesimpulan berdasarkan karakteristik setiap anggota *Cluster* untuk menentukan wilayah yang terdampak banjir tahun 2020 dan 2021.
- Membuat peta spasial hasil dari *Clustering* data tahun 2020 dan 2021 dengan menggunakan ArcGis versi 10.8.

2.2. Validasi Cluster Data Tahun 2020 dan 2021

Cluster yang telah terbentuk perlu untuk dilakukan validasi, tujuannya agar dapat diketahui data tersebut layak dibagi kedalam berapa *cluster*. Untuk validasi *cluster* ada beberapa cara

salah satunya dengan *Davies bouldin index*. Validasi *Clustering* digunakan untuk memutuskan *cluster* terbaik pada dataset yang diujikan (Tempola, dkk., 2020). DBI merupakan salah satu metode evaluasi *cluster* pada suatu metode pengelompokan yang didasarkan pada nilai kohesi dan separasi. Dalam suatu pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap *centroid* dari *cluster* yang diikuti. Sedangkan separasi didasarkan pada jarak antar *centroid* dari *Clusternya* (Hablum, dkk., 2019).

Ada empat tahapan dalam menghitung DBI diantaranya yaitu (Septiani, dkk., 2022):

- a) Menghitung *Sum Of Square Within Cluster* (SSW) ialah keterikatan anggota satu *cluster* atau seberapa mirip antara anggota satu dan dua dan semakin kecil semakin bagus dikarenakan semakin mirip. SSW dihitung untuk mengetahui matrik/kohesi/homogenitas. Kohesi merupakan keterikatan anggota *cluster* dalam satu *cluster*.

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_j) \quad (1)$$

Keterangan:

m_i = jumlah data dalam *Cluster* ke-i

$d(x, c)$ = jarak data ke *centroid*

x_j = data pada *Cluster* ke-j

c_i = *centroid Cluster* ke-i

- b) Menghitung *Sum Of Square between Cluster* (SBB) merupakan jarak antar *cluster* cukup besar sehingga terpisah ke dalam kelompok lain. SSB bertujuan untuk mengetahui separasi/heterogenitas. Separasi merupakan perbedaan antara satu *cluster* dengan *cluster* lainnya.

$$SSB_{i,j} = d(c_i, c_j) \quad (2)$$

Keterangan:

C_i = *cluster* 1

C_j = *cluster* lainnya

$d(C_i, C_j)$ = Jarak antara *centroid* satu dengan lainnya

- c) Menghitung Rasio berfungsi untuk mengetahui seberapa bagus nilai perbandingan *cluster* satu dengan *cluster* lainnya. Jumlah kohesi harus kecil sedangkan jumlah separasi harus lebih besar.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (3)$$

Keterangan:

$R_{i,j}$ = Rasio antar *cluster* ke-i dan ke-j

SSW_i = SSW *cluster* ke-i

SSW_j = SSW *cluster* ke-j

SSB_{ij} = separasi dari *cluster* i dan j

- d) Menghitung DBI (*Davis Bouldin Index*), Faktanya jika hasil dari perhitungan DBI yang diperoleh semakin mendekati nol akan tetapi tidak negatif, maka nilai hasil *Clustering* semakin baik.

$$DBI = \frac{1}{k} \sum_{i=1}^k \text{MAX}_{i \neq j} (R_{i,j}) \quad (4)$$

Keterangan:

k = kluster yang ada

$R_{i,j}$ = rasio antara kluster i dan j

Max = Maksimum $R_{i,j}$

3. Results and Discussion

Proses *Clustering* menggunakan K-Means dalam penelitian ini bertujuan untuk mengelompokkan data berdasarkan karakter yang sama dalam satu *cluster* dan memisahkan data dengan karakter yang berbeda ke *cluster* yang lain. Data yang digunakan dalam perhitungan pada tahap ini adalah data banjir pada tahun 2020 dan 2021. Pada penelitian ini digunakan k= 3,4 dan 5 yang kemudian divalidasi untuk menentukan *cluster* yang optimal.

3.1. Iterasi Pertama

a. Menentukan centroid awal (k=3) untuk tahun 2020

Dalam menentukan nilai *centroid* untuk awal iterasi, nilai awal *centroid* dilakukan secara acak (Wahidiah, 2010). Tiga *centroid* awal dipilih secara acak seperti pada Tabel 1.

Tabel 1. *Centroid* awal k=3 secara acak data 2020

| DATA | Centroid | X1 | X2 | X3 | X4 | X5 |
|------|----------|-----|----|------|-----|----|
| D19 | C1 | 134 | 5 | 11,7 | 0,5 | 0 |
| D20 | C2 | 184 | 5 | 300 | 3 | 0 |
| D26 | C3 | 42 | 0 | 45 | 2 | 0 |

3.2. Menghitung jarak antara titik centroid dan Menentukan titik cluster setiap data tahun 2020

Untuk menghitung jarak semua data maka diperlukan perhitungan jarak untuk setiap data. Dalam perhitungan jarak digunakan rumus *Euclidean Distance*, yaitu:

$$d_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$

Perhitungan jarak menggunakan *Euclidean Distance*, Perhitungan jarak tersebut dilakukan untuk menentukan *centroid 1* (C1), *centroid 2* (C2) dan *centroid 3* (C3) semua data banjir (D1-D46) tahun 2020.

Berdasarkan perhitungan jarak dengan *Euclidean Distance* yang dilakukan selanjutnya ditentukan *cluster* untuk setiap data atau iterasi pertama. melihat nilai minimum jaraknya. Jika jarak minimum berada pada C1 maka data tersebut merupakan anggota *Cluster 1* dan begitupun jika jarak minimum berada pada C2 dan C3 seperti pada Tabel 2.

Tabel 2. Hasil perhitungan jarak dan letak *Cluster data 2020*

| DATA | C1 | C2 | C3 | Minimum | Cluster |
|------|--------|--------|--------|---------|---------|
| D1 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D10 | 84,96 | 328,62 | 45,75 | 45,75 | 3 |
| D11 | 114,41 | 324,55 | 33,36 | 33,36 | 3 |
| D12 | 134,51 | 351,94 | 61,79 | 61,79 | 3 |
| D13 | 604,84 | 574,03 | 691,23 | 574,03 | 2 |
| D14 | 202,11 | 215,52 | 272,84 | 202,11 | 1 |
| D15 | 256,77 | 330,40 | 347,24 | 256,77 | 1 |

| DATA | C1 | C2 | C3 | Minimum | Cluster |
|------|--------|--------|--------|---------|---------|
| D16 | 111,58 | 309,10 | 19,82 | 19,82 | 3 |
| D17 | 134,75 | 330,93 | 46,56 | 46,56 | 3 |
| D18 | 171,26 | 158,44 | 219,77 | 158,44 | 2 |
| D19 | 0,00 | 292,61 | 97,98 | 0,00 | 1 |
| D2 | 110,69 | 340,04 | 48,52 | 48,52 | 3 |
| D20 | 292,61 | 0,00 | 291,92 | 0,00 | 2 |
| D21 | 266,29 | 369,68 | 360,91 | 266,29 | 1 |
| D22 | 327,37 | 408,40 | 421,58 | 327,37 | 1 |
| D23 | 80,84 | 277,79 | 21,59 | 21,59 | 3 |
| D24 | 148,71 | 289,93 | 52,21 | 52,21 | 3 |
| D25 | 149,46 | 315,92 | 245,20 | 149,46 | 1 |
| D26 | 97,98 | 291,92 | 0,00 | 0,00 | 3 |
| D27 | 152,31 | 265,15 | 63,64 | 63,64 | 3 |
| D28 | 99,35 | 317,19 | 26,49 | 26,49 | 3 |
| D29 | 601,17 | 369,70 | 625,64 | 369,70 | 2 |
| D3 | 124,62 | 346,84 | 55,26 | 55,26 | 3 |
| D30 | 133,50 | 250,07 | 55,61 | 55,61 | 3 |
| D31 | 246,04 | 205,77 | 178,59 | 178,59 | 3 |
| D32 | 140,88 | 311,11 | 46,83 | 46,83 | 3 |
| D33 | 49,66 | 315,65 | 62,97 | 49,66 | 1 |
| D34 | 20,45 | 301,98 | 117,02 | 20,45 | 1 |
| D35 | 51,60 | 316,28 | 61,59 | 51,60 | 1 |
| D36 | 99,82 | 335,01 | 45,59 | 45,59 | 3 |
| D37 | 81,01 | 327,01 | 46,62 | 46,62 | 3 |
| D38 | 176,46 | 325,44 | 271,76 | 176,46 | 1 |
| D39 | 45,81 | 300,12 | 143,27 | 45,81 | 1 |
| D4 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D40 | 202,75 | 90,22 | 207,42 | 90,22 | 2 |
| D41 | 97,83 | 334,13 | 45,32 | 45,32 | 3 |
| D42 | 31,67 | 310,28 | 77,45 | 31,67 | 1 |
| D43 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D44 | 24,04 | 305,53 | 100,86 | 24,04 | 1 |
| D45 | 120,67 | 344,87 | 53,04 | 53,04 | 3 |
| D46 | 101,34 | 335,02 | 54,28 | 54,28 | 3 |
| D5 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D6 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D7 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |
| D8 | 132,61 | 350,94 | 60,24 | 60,24 | 3 |
| D9 | 134,60 | 351,98 | 61,59 | 61,59 | 3 |

3.3. Mencari centroid baru

Setelah mendapatkan jumlah anggota setiap *cluster*, selanjutnya mencari *centroid* baru dan melanjutkan iterasi kedua, dengan tahap yang sama dengan iterasi pertama. Dalam menentukan centroid baru dapat dilihat nilai setiap *cluster* pada hasil iterasi pertama Rumus yang digunakan dalam mencari *centroid* baru yaitu rumus:

$$C_k = \frac{1}{n_k} \sum d_i \tag{6}$$

Persamaan (6) digunakan untuk mencari C1 pada X1 hingga X5. Setelah mendapatkan hasil C1 baru selanjutnya dengan cara yang sama dilakukan untuk mencari C2 dan C3 baru, kemudian dilanjutkan ke iterasi kedua.

Tabel 3. Hasil Perhitungan *Centroid* Baru

| <i>Centroid</i> | X1 | X2 | X3 | X4 | X5 |
|-----------------|--------|------|-------|------|------|
| C1 | 231,85 | 3,23 | 13,36 | 1,12 | 1,54 |
| C2 | 347 | 4,8 | 256,4 | 1 | 40 |
| C3 | 17,61 | 0,46 | 26,29 | 0,23 | 3,04 |

3.4. Iterasi kedua

Berdasarkan hasil iterasi kedua apabila terdapat data yang berpindah *cluster* maka dilanjutkan ke iterasi ketiga, jika pada iterasi ketiga masih ada data yang berpindah *cluster* maka dilanjutkan ke iterasi selanjutnya hingga data setiap anggota *cluster* tidak berpindah atau tidak mengalami perubahan. Dalam menentukan hasil untuk iterasi ketiga dan seterusnya makan dilakukan cara yang sama pada iterasi sebelumnya yaitu dengan mencari centroid baru, melakukan perhitungan jarak baru dan menentukan letak *cluster* setiap data. Pada k=3 data tahun 2020 iterasi yang dilakukan sampai dengan iterasi kelima yang menghasilkan data pada setiap *cluster* sudah tidak mengalami perubahan dari iterasi sebelumnya (iterasi keempat).

Dengan cara yang sama dilakukan pada k=4 dan k=5 untuk data tahun 2020 dan untuk menentukan k=3, 4 dan 5 data tahun 2021.

3.5. Hasil validasi data tahun 2020 dan 2021

Untuk menganalisis hasil dari *Clustering* menggunakan metode *K-Means* yang telah didapatkan, dibutuhkan uji validasi untuk menghitung keakuratan setiap *cluster* dan mendapatkan jumlah *cluster* terbaik. Pada penelitian ini digunakan metode *Davies-Bouldin Index* (DBI) dalam mengvalidasi *cluster*. Tabel 4 – 8 menunjukkan hasil DBI yang diperoleh.

Tabel 4. Nilai DBI untuk K=3 Data Tahun 2020

| R | 1 | 2 | 3 | R MAX | DBI |
|---|------|------|------|-------|------|
| 1 | 0 | 1,04 | 0,68 | 1,04 | 0,92 |
| 2 | 1,04 | 0 | 0,50 | 1,04 | |
| 3 | 0,68 | 0,50 | 0 | 0,68 | |

Tabel 5. Nilai DBI untuk K=4 Data Tahun 2020

| R | 1 | 2 | 3 | 4 | R MAX | DBI |
|---|------|------|------|------|-------|------|
| 1 | 0 | 0,2 | 0,26 | 0,81 | 0,81 | 0,73 |
| 2 | 0,2 | 0 | 0,66 | 0,35 | 0,66 | |
| 3 | 0,26 | 0,66 | 0 | 0,34 | 0,66 | |
| 4 | 0,81 | 0,35 | 0,34 | 0 | 0,81 | |

Tabel 6. Nilai DBI untuk K=5 Data Tahun 2020

| R | 1 | 2 | 3 | 4 | 5 | R MAX | DBI |
|---|------|------|------|------|------|-------|------|
| 1 | 0,00 | 0,37 | 0,52 | 0,52 | 0,67 | 0,67 | |
| 2 | 0,37 | 0,00 | 0,52 | 0,95 | 0,42 | 0,95 | |
| 3 | 0,52 | 0,52 | 0,00 | 0,62 | 0,49 | 0,62 | 0,79 |
| 4 | 0,52 | 0,95 | 0,62 | 0,00 | 0,74 | 0,95 | |
| 5 | 0,67 | 0,42 | 0,49 | 0,74 | 0,00 | 0,74 | |

Tabel 7. Nilai DBI Data Tahun 2020

| 3 Cluster | 4 Cluster | 5 Cluster |
|-----------|-----------|-----------|
| 0,92 | 0,73 | 0,79 |

Dengan cara yang sama dalam menentukan nilai DBI pada data tahun 2020 juga dilakukan pada data tahun 2021, dapat dilihat pada tabel 8.

Tabel 8. Nilai DBI Data Tahun 2021

| DBI k=3 | DBI k=4 | DBI k=5 |
|---------|---------|---------|
| 0,75 | 0,44 | 0,71 |

Dari ketiga hasil perhitungan DBI tersebut, dihasilkan DBI minimum pada k=4 untuk data tahun 2020 dan 2021 sehingga digunakan 4 *cluster* dalam *Clustering* data banjir kabupaten bone tahun 2020 dan 2021 dengan DBI yaitu 0,73 dan 0,44. Dalam pengujian DBI, apabila nilai yang diberikan mendekati 0 atau non-negatif ≥ 0 , maka hasil *cluster* yang diberikan semakin baik (Ramadhani, dkk., 2022).

Tabel 9. Nama Desa Anggota Setiap *Cluster* Data Tahun 2020

| Cluster | Anggota Cluster |
|---------|---|
| 1 | Bulu Tempe, Pacubbe, Panyiw, Kading, Cakke Bone, Kajuara, Carigading, Macompe, Polewali, Kading, Watu, Massila, Pinceng Pute, Labissa, Unyi, Patanga, Pacubbe, Labotto, Kanco, Itterung, Carigading, Panyiw, Labonge, Awang Cenrana, Padacenga, Tocina, Lamuru, Matajang, Ajanglase, Ujung, Nagauleng, Pattiro, Watu, Sijelling |
| 2 | Pompanua Riattang, Welado, Uloe, Tawaroe, Solo, Ujung Tanah, Kampoti |
| 3 | Pompanua, Lea |
| 4 | Pakkasalo, Pallae, Waji |

Tabel 10. Deskripsi Tiap *Cluster* Data Tahun 2020

| Cluster | X1 | X2 | X3 | X4 | X5 |
|---------|--------|------|--------|-------|-------|
| 1 | 40 | 0,52 | 15,66 | 0,202 | 2,059 |
| 2 | 340,29 | 6 | 44,58 | 2,29 | 0 |
| 3 | 575,5 | 7 | 310 | 0 | 100 |
| 4 | 116,33 | 1,67 | 242,33 | 1 | 11,67 |

Desa pada *Cluster* 1 memiliki jumlah keluarga yang terdampak (X1) lebih rendah dibandingkan kelaster lainnya dengan jumlah rata 40 KK selain itu luas sawah yang mengalami kerusakan (X3) juga paling rendah dibandingkan dengan *cluster* lainnya.

Desa pada *Cluster* 2 memiliki jumlah keluarga yang terdampak banjir (X1) tertinggi kedua setelah *cluster* 1 dan tertinggi kedua luas sawah yang mengalami kerusakan, tetapi untuk kebun dan kolam ikan (X5) yang mengalami kerusakan yaitu 0 Ha lebih rendah dibandingkan dengan *cluster* lainnya.

Desa pada *Cluster 3* memiliki jumlah keluarga yang terdampak (X1), banyak unit fasilitas yang rusak (X2) dan luas sawah yang rusak (X3) lebih tinggi dari semua *cluster*. Sedangkan untuk panjang jalan yang rusak (X4) dengan jumlah 0 km lebih rendah dari *cluster* lainnya.

Desa pada *Cluster 4* memiliki jumlah keluarga yang terdampak (X1) dan banyak unit yang rusak (X2) dibandingkan *cluster 1*. Sedangkan untuk Luas sawah (X3) dan Luas Kolam (X4) serta panjang jalan (X4) yang mengalami kerusakan tertinggi kedua dibandingkan *cluster 1*.

Tabel 11. Nama Desa Anggota Setiap *Cluster* Data Tahun 2021

| <i>Cluster</i> | Anggota <i>Cluster</i> |
|----------------|--|
| 1 | Mallusetasi, Pakkasalo |
| 2 | Nagauleng, Pattiro Sompe, Mattiro Walie, Polewali, Pasaka, Bulie, Pasempe, Patimpa, Macompe, Kading, Cakke Bone, Carigading, Wallinrang, Cinennung, Pompanua, Parippung, Bulie, Sengeng Palie, Pallime, Welado, Matajang |
| 3 | Kampoti, Unyi |
| 4 | Tocina |

Tabel 12. Deskripsi *Cluster* Data Tahun 2021

| <i>Cluster</i> | X1 | X2 | X3 | X4 | X5 |
|----------------|--------|------|--------|------|--------|
| 1 | 207,67 | 2,67 | 266,67 | 0,17 | 150,00 |
| 2 | 7,64 | 0,36 | 1,45 | 0,00 | 1,36 |
| 3 | 240,00 | 0,00 | 8,33 | 0,12 | 0,00 |
| 4 | 175,00 | 1,00 | 0,00 | 0,00 | 600,00 |

Desa pada *Cluster 1* memiliki keluarga yang terdampak (X1) dan luas kolam (X5) tertinggi kedua sedangkan untuk banyak unit (X2), luas sawah (X3) dan panjang jalan (X4) yang mengalami kerusakan lebih tinggi dibandingkan *cluster* lainnya.

Desa pada *Cluster 2* hampir semua dampak ditimbulkan lebih rendah dari *cluster* lainnya.

Desa pada *Cluster 3* memiliki keluarga yang terdampak (X1) tertinggi dari yang lainnya akan tetapi untuk banyak unit (X2) dan luas kebun (X5) dengan jumlah 0 lebih rendah dari *cluster* lainnya.

Desa pada *Cluster 4* 3 memiliki keluarga yang terdampak (X1) lebih tinggi dari *cluster 2* tetapi untuk luas kebun yang mengalami kerusakan (X5) lebih tinggi dibandingkan yang lainnya. Sedangkan untuk luas sawah (X3) dan panjang jalan (X4) yang mengalami kerusakan lebih rendah dibandingkan yang lainnya yaitu 0.

Variasi spasial data tahun 2020 dan 2021

Berdasarkan data banjir tahun 2020 dan 2021 yang didapatkan di kantor BPBD Kab.Bone dengan jumlah data banjir tahun 2020 sebanyak 46 dan tahun 2021 sebanyak 29 dengan total data sebanyak 75 data. dari ke-75 data tersebut ada 13 desa yang terdampak banjir pada tahun 2020 dan 2021. Dapat dilihat pada Gambar 1 dan 2 untuk Warna hijau pada peta menandakan bahwa wilayah tersebut tidak terdampak banjir sedangkan warna lainnya menandakan bahwa daerah tersebut terdampak banjir.

4. Conclusion

Hasil K-Means Clustering pada data tahun 2020 dan 2021 dihasilkan cluster desa berdasarkan kemiripannya. Dari hasil algoritma tersebut, didapatkan 4 cluster pada tahun 2020 dan 2021 serta anggota cluster-nya. Hasil cluster data tahun 2020 dihasilkan 34 anggota cluster 1, 7 anggota cluster kedua, 2 anggota cluster ketiga dan 3 anggota cluster keempat. Kemudian untuk data tahun 2021 dihasilkan 3 anggota cluster pertama, 22 anggota cluster kedua, 3 anggota cluster ketiga dan 1 anggota cluster keempat. Data tahun 2020 dan 2021 terdapat 13 desa yang terdampak banjir pada tahun 2020 dan mengalami peningkatan

ataupun penurunan untuk keluarga yang terdampak dan kerusakan yang dihasilkan pada tahun 2021.

References

- Aditya. (2021). Banjir Melanda 9 Daerah di Sulawesi Selatan, 3.206 Warga Mengungsi. *Tempo.Co*. <https://nasional.tempo.co/read/1537190/banjir-melanda-9-daerah-di-sulawesi-selatan-3-206-warga-mengungsi>. Diakses pada 2 Desember 2022.
- Alfina, T., & Santosa, B. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS). *Analisa PerbandinganMetode Hierarchical Clustering, K-Means Dan Gabungan Keduanya Dalam Cluster Data*, 1(1). 1–5.
- Ali, M., & Trisutomo, S. (2017). Pemetaan Daerah Rawan Banjir Berbasis Sistem Informasi Geografis (GIS) di Pesisir Danau Tempe Kabupaten Wajo. *Losari: Jurnal Arsitektur Kota Dan Pemukiman*, 1(1). 37–42.
- Anasiru, R. H. (2016). Analisis Spasial dalam Klasifikasi Lahan Kritis di Kawasan Sub-Das Lange Gorontalo Spatial Analysis in the Classification of Critical Land in the Sub-Basin of Lange Gorontalo. *Jurnal Informatika Pertanian*, 25(2). 261–272.
- BPS Bone. (2020). Daerah Terdampak Banjir di Kabupaten Bone. <http://bonekab.bps.go.id>. Diakses pada 20 Desember 2022.
- BPS Bone. (2021). Daerah Terdampak Banjir di Kabupaten Bone. <http://bonekab.bps.go.id>. Diakses pada 20 Desember 2022.
- Faiza, N. A. R. (2019). *Cash Waqf Linked Sukuk sebagai Pembiayaan Pemulihan Bencana Alam di Indonesia*. Thesis. UIN Sunan Ampel, Surabaya.
- Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi. *Computer Engineering, Science and System Journal*, 3(2). 173.
- Hablum, rofika julianti, Khairan, A., & Rosihan. (2019). Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means. *JIKO (Jurnal Informatika Dan Komputer)*, 2(1). 26–33.
- Irina, F. (2017). *Penelitian Terapan (Applied Research)*. <https://openlibrary.telkomuniversity.ac.id/pustaka/143081/metode-penelitian-terapan.html>. Diakses pada 5 Desember 2022
- Ramadhani, S., Azzahra, D., & Z, T. (2022). Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student's Thesis. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 13(1). 24–33. <https://doi.org/10.31849/digitalzone.v13i1.9292>
- Sanusi, W., Zaky, A., & Afni, B. N. (2020). Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-faktor Penyebab Gizi Buruk. *Journal of Mathematics, Computations, and Statistics*, 2(1). 47–56. <https://doi.org/10.35580/jmathcos.v2i1.12458>
- Septiani, I. W., Fauzan, A. C., & Huda, M. M. (2022). Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 3(4). 556–567. <https://doi.org/10.30865/json.v3i4.4055>
- Syarif, R., Furqon, M. T., & Adinugroho, S. (2018). Perbandingan Algoritme K-Means Dengan Algoritme Fuzzy C Means (FCM) Dalam Clustering Moda Transportasi Berbasis GPS. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Branjijaya*, 2(10), 4107–4115. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/2852>
- Tempola, F., Muhammad, M., & Mubarak, A. (2020). Penggunaan Internet Dikalangan Siswa SD di Kota Ternate: Suatu Survey, Penerapan Algoritma Clustering dan Validasi DBI. *Jurnal Teknologi*

Informasi Dan Ilmu Komputer, 7(6). 1153-1163. <https://doi.org/10.25126/jtiik.2020722370>

Wahidah, N. (2010). Clustering Menggunakan K-Means Algorithm. *Jurnal Transformatika*, 8(1). 33-42. <https://doi.org/10.26623/transformatika.v8i1.45>

