



Analysis Prediksi Wilayah Rawan Banjir dengan Algoritma K-Means

Muhammad Makmun Effendi*, Inka, Arif Siswandi

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi

Jl. Inspeksi Kalimalang No.9,Cibatu, Rumbai, Kabupaten Bekasi, Jawa Barat, Indonesia

Email: ^{1,*}effendiyana@pelitabangsa.ac.id ²inkamanik21@gmail.com, ³arif.siswandi@pelitabangsa.ac.id

Submitted: 07/01/2024; Accepted: 31/01/2024; Published: 31/01/2024

Abstrak—Seiring dengan tingginya jumlah curah hujan pada Kabupaten Bekasi-Jawa Barat, banjir sudah mulai menggenangi beberapa wilayah kabupaten Bekasi, salah satu penyebabnya adalah faktor tingginya curah hujan. Menurut (Badan Penanggulangan Bencana Daerah) BPBD titik banjir terbanyak berada di wilayah Bekasi, sehingga menyebabkan beberapa aktivitas masyarakat sekitarnya terganggu, terhambatnya transportasi, dan juga timbulnya penyakit masalah seperti penyakit kulit, diare, dan sebagainya. Permasalahan banjir merupakan tanggung jawab bersama yang memerlukan solusi juga peran teknologi untuk membantu mempermudah pemberian informasi kepada masyarakat mengenai daerah rawan banjir di wilayah Bekasi. Salah satu teknik yang dapat digunakan adalah menggunakan Algoritma K-Means Clustering untuk mengelompokkan wilayah rawan banjir. Dataset banjir tersebut diolah menggunakan aplikasi RapidMiner, untuk dataset yang diambil untuk melakukan analisa tersebut dari bulan Januari hingga Desember 2022 sebanyak 24 data daerah yang terkena banjir dari berbagai kecamatan dan desa di kota Bekasi. Hasil penelitian menghasilkan 3 cluster yaitu, kategori banjir tinggi, banjir sedang, dan banjir rendah mendapatkan nilai davies bouldin index sebesar -0,452.

Kata Kunci: Bekasi; Banjir; K-Means Clustering; BPBD; RapidMiner

Abstract—Along with the high amount of rainfall in Bekasi -West Java, floods have started to inundate several areas of Bekasi, one of the causes is the high rainfall factor. According to (Regional Disaster Management Agency) BPBD, the most flood points are in the Bekasi area, causing several activities of the surrounding community to be disrupted, transportation hampered, and also the emergence of disease problems such as skin diseases, diarrhea, and so on. The problem of flooding is a shared responsibility that requires a solution. also the role of technology to help facilitate the provision of information to the public regarding flood-prone areas in the Bekasi area. One technique that can be used is using the K-Means Clustering Algorithm to group flood-prone areas. The flood dataset was processed using the RapidMiner application, for the dataset taken to carry out this analysis from January to December 2022, there were 24 data from areas affected by flooding from various sub-districts and villages in the city of Bekasi. The results of the research produced 3 clusters, namely, the high flood, medium flood and low flood categories, which received a Davies Bouldin index value of -0.452.

Keywords: Bekasi; Flood; K-Means Clustering; BPBD; Rapid Miner

1. PENDAHULUAN

Anomali iklim adalah perubahan musiman dari rata-rata normal yang terjadi akibat pemanasan global, yakni peningkatan suhu bumi yang menyebabkan penguapan dan kenaikan curah hujan [1]. Salah satu bentuk anomali iklim adalah kondisi La Nina, di mana bagian tengah Samudera Pasifik mengalami pendinginan di bawah kondisi normalnya. Dampak dari La Nina ini meliputi hujan lebat yang dapat menyebabkan bencana banjir. Peningkatan curah hujan yang berlebihan akibat perubahan iklim ini memberikan kontribusi pada kejadian banjir yang serius [2].

Banjir sering terjadi di daerah perkotaan dan merupakan salah satu bencana alam hidrometeorologi yang paling umum. Beberapa faktor yang mempengaruhi terjadinya banjir meliputi curah hujan tinggi, elevasi yang datar, dan kondisi lahan yang landai. Selain itu, perilaku masyarakat juga turut berperan, seperti pembuangan sampah sembarangan dan penebangan pohon tanpa perhatian terhadap lingkungan. Dampak banjir pada suatu wilayah bisa sangat merugikan, mencakup kerusakan lingkungan, kerugian materi yang besar, dan hilangnya nyawa [3].

Peningkatan jumlah penduduk juga menjadi penyebab banjir, dengan meningkatnya jumlah penduduk maka kebutuhan lahan juga semakin meningkat akhirnya terjadi pengalihan fungsi lahan yang sebelumnya lahan pertanian diubah menjadi pemukiman yang menyebabkan berkurangnya tutupan lahan alami yang berfungsi sebagai daerah resapan. Menurut (Badan Penanggulangan Bencana Daerah) BPBD titik banjir terbanyak berada di wilayah Bekasi, hal ini mengakibatkan beberapa aktivitas pekerjaan masyarakat terganggu, terhambatnya transportasi, juga masalah kesehatan seperti penyakit kulit, diare, dan lain-lain. Permasalahan banjir sendiri merupakan tanggung jawab bersama yang memerlukan solusi bukan saja dari pemerintah saja namun juga dari seluruh masyarakat khususnya di wilayah Bekasi. Untuk itu perlu meningkatkan kesiagaan masyarakat agar lebih berhati-hati terhadap bencana banjir dan juga dalam hal ini sangat diperlukan peran teknologi yaitu data yang akurat dan tepat mengenai daerah-daerah yang rawan banjir untuk membantu mempermudah pemberian informasi kepada masyarakat mengenai daerah rawan banjir di wilayah Bekasi [4].

Penelitian terkait prediksi banjir telah dengan menerapkan proses klasifikasi menggunakan algoritma deep learning H2O untuk memprediksi menggunakan data curah hujan menunjukkan bahwa akurasi yang baik ialah senilai 90.54% dan RMSE senilai 0.087 menggunakan jumlah epoch 30 dengan fungsi aktivasi rectifier [5]. penelitian selanjutnya dengan menggunakan algoritma K-Mean dan K-Medoids pengelompokan rawan banjir adalah dengan algoritma K-Means menghasilkan 0.218 dengan kluster optimal adalah K= 3 sedangkan K-Medoids

nilai validasinya 0.525 dengan kluster optimal adalah $K=6$ [6]. Penelitian berikutnya dalam melakukan prediksi banjir dengan membandingkan tiga algoritma yaitu Decision Tree, rata-rata yang diperoleh masing-masing akurasi, precision, recall, dan f1-score, yakni sebesar 99,05%, 97,91%, Random Forest, dan Naïve Bayes maka hasil yang didapat berdasarkan rasio dari pembagian data yang berbeda, yakni 8:2, 7:3, dan 6:4. menunjukkan algoritma dengan akurasi dan performa yang paling baik dimiliki oleh algoritma Random Forest jika dibandingkan dengan metode Decision Tree dan naïve bayes. Nilai 99,18%, 98%, kemudian rata-rata nilai waktu komputasi sebesar 0,2561 detik [7]. Penelitian yang lainnya mengenai deteksi bencana banjir dengan menggunakan metode Naïve Bayes dengan tools Rapidminer menghasilkan akurasi sebesar 76.73% [8]

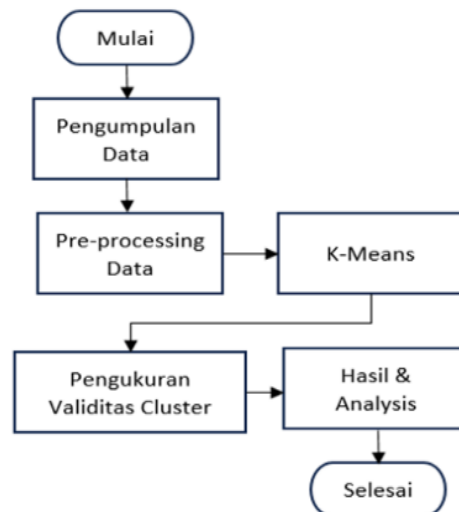
Data Mining adalah proses ekstraksi informasi dari sejumlah besar data untuk menemukan pengetahuan yang tersembunyi dan bisa digunakan secara langsung. Proses Data Mining melibatkan beberapa metode analisis data, seperti Clustering, Association, dan Klasifikasi, dan teknik lainnya. Clustering adalah cara yang efektif untuk menganalisis data dengan cara yang lebih eksploratif dan dapat digunakan di berbagai bidang. Metode Clustering dibagi menjadi tiga kategori utama, yaitu partisi hierarki, berbasis grid, dan berbasis model [9].

Salah satu teknik yang dapat digunakan adalah Algoritma K-Means Clustering dikarenakan klasifikasi merupakan sebuah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep dan kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui. untuk mengelompokan wilayah rawan banjir. menjadi objek clustering. Metode K-Means Clustering dipilih karena mudah diimplementasikan dan diadaptasi serta umum digunakan dalam penyelesaian studi [10]

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam penelitian ini penulis menggunakan algoritma K-Means Clustering. Algoritma K-means merupakan salah satu algoritma dengan partitional, karena K-Means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya. Algoritma K-Means menggunakan proses secara berulang-ulang dan salah satu metode data clustering non-hirarki yang berusaha mempartisi data yang ada ke dalam satu bentuk atau lebih cluster [11]. Data yang digunakan merupakan data sekunder yang didapatkan dari website resmi Adapun data yang diambil adalah kecamatan yang ada di bekasi dengan atribut faktor penyebab banjir yaitu jumlah curah hujan pada bulan Januari hingga Oktober 2022 yang diambil oleh penulis adalah sebanyak 50 data kecamatan di kota Bekasi yang paling sering terkena banjir berekstensi xlsx [12]. Output yang diharapkan adalah menghasilkan 3 cluster yaitu kategori banjir tinggi (C0), banjir sedang (C1), dan banjir rendah (C2). Variabel atau atribut digunakan dalam pengelompokan dataset banjir ini terdiri dari jumlah titik banjir, bulan yaitu waktu terjadinya hujan, serta curah hujan [13]. Gambar 1 berikut merupakan tahapan yang dilakukan dalam penelitian:



Gambar 1. Metodologi Penelitian

1. Tahap pengumpulan data, yaitu dilakukan pencarian terhadap dataset banjir pada tahun 2020 dan bulan Januari hingga Oktober 2022 yang didapat dari website <https://danta.bekasikota.go.id/user/ViewDataset?q=jumlah-bencana-banjir> dan <https://danta-admin.bekasikota.go.id/hu/dataset/jumlah-bencana-banjir/resource/f8ff6640-6de8-4c1f-be33-9487d582d315>.
2. Tahap Pre-Processing Data. Dalam tahap ini dilakukan analisa tentang algoritma yang cocok digunakan untuk dataset yang ada yaitu k-means.



3. Tahap implementasi yaitu, dilakukan pengujian terhadap dataset yang didapat dengan algoritma k-means dengan cara menghitung centroid pada setiap satuan data.
4. Tahap Hasil dan Analisis. Pada tahap ini semua hasil hitung dilampirkan kedalam laporan serta grafik yang dihasilkan ditampilkan sebagai representasi tiap data.

2.2 Pengumpulan Data

Data yang digunakan merupakan data sekunder yang didapatkan dari website resmi Adapun data yang diambil adalah kecamatan yang ada di bekasi dengan atribut faktor penyebab banjir yaitu jumlah curah hujan pada bulan Januari hingga Oktober 2022 yang diambil oleh penulis adalah sebanyak 50 data kecamatan di Kota Bekasi yang paling sering terkena banjir berekstensi xlsx. Output yang diharapkan adalah menghasilkan 3 cluster yaitu kategori banjir tinggi (C0), banjir sedang (C1), dan banjir rendah (C2) Variabel atau atribut digunakan dalam pengelompokan dataset banjir ini terdiri dari jumlah titik banjir, bulan yaitu waktu terjadinya hujan, serta curah hujan [14].

Tabel 1. Dataset Banjir

Lokasi Kejadian (Kecamatan)	Jumlah Titik Banjir	Bulan	Curah Hujan Rata-Rata
Cabangbungin	1	Januari	116 mm
Jatiasih	1	Februari	127 mm
Bekasi Timur	3	Februari	127 mm
Bekasi Selatan	2	Februari	127 mm
Bekasi Utara	2	Februari	127 mm
Pondok Gede	4	April	75 mm
Bekasi Barat	5	April	75 mm
Bekasi Timur	1	April	75 mm
Jatiasih	1	April	75 mm
Pondok Gede	4	Mei	48 mm
Pondok Melati	1	Mei	48 mm
Bekasi Timur	5	Mei	48 mm
Jatiasih	1	Mei	48 mm
Jatiasih	1	Juni	50 mm
Pondok Gede	1	Juni	50 mm
Rawalumbu	1	Juni	50 mm
Jatiasih	4	Juli	43 mm
Bekasi Timur	4	Juli	43 mm
Bekasi Selatan	3	Juli	43 mm
Bekasi Utara	1	Juli	43 mm
Pondok Gede	4	Oktober	34 mm
Jatiasih	1	Oktober	34 mm
Rawalumbu	2	Oktober	34 mm
Bekasi Barat	1	Oktober	34 mm

2.3 Preprocessing Data

Preprocessing data atau proses mencari data mentah dengan menggunakan dataset Banjir.xlsx yang bersumber dari data Bekasi. Setelah itu melakukan proses data cleaning yaitu proses pembersihan data yaitu merubah data, tujuannya adalah untuk mengolah data serta memperbaiki kesalahan-kesalahan yang ada pada dataset dengan menggunakan algoritma K-Means Clustering dan algoritma ini hanya bisa diolah menggunakan angka sub title 2 ini bisa juga berisi metode penyelesaian masalah, serta tahapan tahapan dari metode tersebut. Dalam naskah, nomor kutipan secara berurutan dalam tanda kurung siku, juga tabel angka dan angka secara berurutan seperti yang ditunjukkan pada tabel 1 dan gambar 1.

Tabel 2. Perubahan Data Menjadi Numerik

Lokasi Kejadian (Kecamatan)	Jumlah Titik Banjir	Bulan	Curah Hujan Rata-Rata
Cabangbungin	1	1	116
Jatiasih	1	2	127
Bekasi Timur	3	2	127
Bekasi Selatan	2	2	127
Bekasi Utara	2	2	127
Pondok Gede	4	4	75
Bekasi Barat	5	4	75
Bekasi Timur	1	4	75
Jatiasih	1	4	75



Lokasi Kejadian (Kecamatan)	Jumlah Titik Banjir	Bulan	Curah Hujan Rata-Rata
Pondok Gede	4	5	48
Pondok Melati	1	5	48
Bekasi Timur	5	5	48
Jatiasih	1	5	48
Jatiasih	1	6	50
Pondok Gede	1	6	50
Rawalumbu	1	6	50
Jatiasih	4	7	43
Bekasi Timur	4	7	43
Bekasi Selatan	3	7	43
Bekasi Utara	1	7	43
Pondok Gede	4	10	34
Jatiasih	1	10	34
Rawalumbu	2	10	34
Bekasi Barat	1	10	34

2.4 Penerapan Algoritma

Dalam penelitian ini penulis menggunakan metode algoritma K-Means Clustering yaitu salah satu metode data clustering non-hirarki yang berusaha mempartisi data yang ada ke dalam satu bentuk atau lebih cluster. Untuk mengelompokkan wilayah rawan banjir. Pada tahapan ini penggunaan algoritma K-Means digunakan untuk mendapatkan hasil cluster data yang akan dibagi menjadi 3 cluster [15].

3. HASIL DAN PEMBAHASAN

Data yang telah dikumpulkan kemudian dilakukan tahap pengelompokkan atau clustering menggunakan algoritma k-means. Berikut ini adalah proses yang dilakukan untuk mengelompokkan wilayah rawan banjir di Kota Bekasi dengan asumsi bahwa parameter input adalah jumlah dataset sebanyak n data dan jumlah inialisasi centroid k=3 sesuai dengan penelitian [16]

Data yang diambil untuk penelitian berjumlah 50 sesuai data yang tersedia pada website untuk dijadikan contoh penerapan algoritma k-means. Percobaan dilakukan dengan menggunakan parameter-parameter berikut :

- Menentukan jumlah cluster

Jumlah cluster : 3
 Jumlah data : 50
 Jumlah atribut : 3

- Menentukan jumlah centroid

Centroid atau pusat awal cluster ditentukan secara acak, dimana nilai cluster 0 diambil pada baris ke-7, nilai cluster 1 diambil pada baris ke- 14, nilai cluster 2 diambil pada baris ke-21.

Tabel 3. Titik Awal Pusat Cluster

Cluster	Daerah	Jumlah Titik Banjir	Bulan	Curah hujan rata rata
C0	Bekasi Barat	5	4	75
C1	Jatiasih	1	6	50
C2	Pondok Gede	4	10	34

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut:

- Tentukan jumlah cluster
- Alokasikan data ke dalam cluster secara random
- Hitung centroid/ rata-rata dari data yang ada di masing-masing cluster
- Alokasikan masing-masing data ke centroid/ rata-rata terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut [17]:

$$D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2} \dots \tag{1}$$

Dimana :

D(i,j) = Jarak data ke i ke pusat cluster j

Xki = Data ke i pada atribut data ke k

Xkj = Titik pusat ke j pada atribut ke k

- Menentukan centroid

Perhitungan jarak dari data ke-1 terhadap pusat cluster :

$$C0 = \sqrt{(1-5)^2} + \sqrt{(1-4)^2} + \sqrt{(116 - 75)^2} = 41,3$$

$$C1 = \sqrt{(1-1)^2} + \sqrt{(1-6)^2} + \sqrt{(116 - 50)^2} = 66,1$$

$$C2 = \sqrt{(1- 4)^2} + \sqrt{(1- 10)^2} + \sqrt{(116 - 34)^2} = 82,5$$

Perhitungan jarak dari data ke-2 terhadap pusat cluster :

$$C0 = \sqrt{(1-5)^2} + \sqrt{(2-4)^2} + \sqrt{(127 - 75)^2} = 52,1$$

$$C1 = \sqrt{(1-1)^2} + \sqrt{(2-6)^2} + \sqrt{(127 - 50)^2} = 77,1$$

$$C2 = \sqrt{(1- 4)^2} + \sqrt{(2-10)^2} + \sqrt{(127- 34)^2} = 93,3$$

Perhitungan jarak dari data ke-3 terhadap pusat cluster :

$$C0 = \sqrt{(3-5)^2} + \sqrt{(2-4)^2} + \sqrt{(127 - 75)^2} = 52$$

$$C1 = \sqrt{(3-1)^2} + \sqrt{(2-6)^2} + \sqrt{(127 - 50)^2} = 52,1$$

$$C2 = \sqrt{(3- 4)^2} + \sqrt{(2-10)^2} + \sqrt{(127- 34)^2} = 93,3$$

3.1 Hasil Analisa Data

Sebagai alat pengujian dataset untuk penerapan algoritma pada penelitian ini penulis menggunakan tool rapidminer 9.0 untuk mendapatkan clustering otomatis dari setiap algoritma [18]. Akurasi algoritma akan diukur menggunakan Davies Bouldin Index [19]. Pada tahap pengujian ini dilakukan 3 proses yaitu Read Excel, Clustering, Performance yaitu dilakukan operasi pencarian nilai davies bouldin index. Pada tahap pengujian ini dilakukan 3 proses yaitu :

1. Read excel yaitu dilakukan operasi penginputan dataset berupa file berekstensi .xls dataset banjir, seperti terlihat pada gambar 2 berikut.

Role	Name	Type	Statistics	Range	Missings
id	Lokasi Kejadian (Kecamatan)	polynomial	mode = Jatiasih (8), least = Cab	Cabangbungin (1), Jatiasih (8),	0
cluster	cluster	nominal	mode = cluster_1 (42), least = c	cluster_1 (42), cluster_2 (2), clu:	0
regular	Jumlah Tikik Banjir	integer	avg = 5.180 +/- 4.448	[1.000 ; 17.000]	0
regular	Bulan	integer	avg = 5.280 +/- 2.900	[1.000 ; 12.000]	0
regular	Curah Hujan rata-rata	integer	avg = 184.400 +/- 322.700	[30.000 ; 1898.000]	0

Gambar 2. Example Set Statistics

2. Clustering yaitu pada tahap ini dilakukan operasi clustering atau pengelompokan sebagai algoritma yang digunakan pada penelitian ini.
3. Performance yaitu dilakukan operasi pencarian nilai davies bouldin index.

Pada tahapan ini ditampilkan hasil statistik dari data yang sudah di uji. Pada tabel terdapat 3 atribut yaitu min sebagai nilai terendah pada tabel dataset, max sebagai nilai tertinggi pada tabel dataset dan average sebagai nilai rata-rata dari penjumlahan tabel dataset tersebut. Gambar 3 berikut merupakan hasil pengujian dan memperoleh performance vector.

```

PerformanceVector:
Avg. within centroid distance: -9325.853
Avg. within centroid distance_cluster_0: -11552.806
Avg. within centroid distance_cluster_1: -1660.174
Avg. within centroid distance_cluster_2: -163624.250
Davies Bouldin: -0.452
    
```

Gambar 3. Performance Vector

Semakin kecil Davies Bouldin Index maka semakin baik cluster yang diperoleh dari pengelompokan menggunakan metode algoritma k-means clustering yaitu menunjukan nilai -0,452 angka tersebut memiliki hasil yang cukup baik karena mendekati angka 0 [16] seperti terlihat pada tabel 4 berikut.

Tabel 4. Hasil Operasi Performance Vector

Performance Vector	Value
Avg. within centroid distance	-9325.853
Avg. within centroid distance cluster 0	-11552.806
Avg. within centroid distance cluster 1	-1660.174
Avg. within centroid distance cluster 2	-16362



Performance Vector	Value
Davies Bouldin	-0.452

Tabel 5 berikut merupakan hasil clustering baik dari hasil pengujian maupun dengan menerapkan Metode K-Means

Tabel 5. Hasil Clustering

Row No	Lokasi Kejadian (Kecamatan)	Cluster	Jumlah Titik Banjir	Bulan	Curah Hujan Rata-Rata
39	Pondok Melati	Cluster_0	13	5	655
40	Jatiasih	Cluster_0	7	6	655
41	Bantar Gebang	Cluster_0	13	7	433
46	Bekasi Barat	Cluster_0	6	10	574
47	Medan Satria	Cluster_0	4	11	422
48	Bekasi Utara	Cluster_0	5	12	407
1	Cabang Bungin	Cluster_1	1	1	116
2	Jatiasih	Cluster_1	1	2	127
3	Bekasi Timur	Cluster_1	3	2	127
4	Bekasi Selatan	Cluster_1	2	2	127
5	Bekasi Utara	Cluster_1	2	2	127
6	Pondok Gede	Cluster_1	4	4	75
7	Bekasi Barat	Cluster_1	5	4	75
8	Bekasi Timur	Cluster_1	1	4	75
9	Jatiasih	Cluster_1	1	4	75
10	Pondok Gede	Cluster_1	4	5	48
11	Pondok Melati	Cluster_1	1	5	48
12	Bekasi Timur	Cluster_1	5	5	48
13	Jatiasih	Cluster_1	1	5	48
14	Jatiasih	Cluster_1	1	6	50
15	Pondok Gede	Cluster_1	1	6	50
16	Rawalumbu	Cluster_1	1	6	50
17	Jatiasih	Cluster_1	4	7	43
18	Bekasi Timur	Cluster_1	4	7	43
19	Bekasi Selatan	Cluster_1	3	7	43
20	Bekasi Utara	Cluster_1	1	7	43

Pada tabel 5 hasil analisa dataset banjir diatas dapat disimpulkan bahwa hasil Cluster 0 memiliki 6 anggota, Cluster 1 memiliki 42 anggota, Cluster 2 memiliki 2 anggota

4. KESIMPULAN

Dari hasil penelitian yang dilakukan oleh penulis, dapat diambil kesimpulan bahwa penerapan algoritma k-means mengelompokkan dataset banjir pada tahun 2022 menghasilkan tiga cluster yaitu, kategori banjir tinggi, banjir sedang, dan banjir rendah. Hasil pengujian dataset banjir mendapatkan nilai davies bouldin index sebesar -0,452 yang memiliki arti semakin kecil nilainya maka semakin baik cluster yang diperoleh.

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada teman teman yang telah memberikan kontribusi pikiran dan tenaga untuk menyelesaikan penelitian ini dan juga kepada Universitas Pelita Bangsa yang telah memberikan sarana dan prasarana serta supportnya untuk peneliti bisa berkarya, semoga penelitian ini bisa bermanfaat untuk semuanya.

REFERENCES

- [1] W. N. Harahap, B. Yuniasih, and S. Gunawan, "Dampak La Nina 2021-2022 terhadap Peningkatan Curah Hujan," vol. 7, no. 1, pp. 26–32, 2023.
- [2] D. Rogi, J. E. X. Rogi, and S. Runtuwuwu, "Analisis Anomali Iklim Dalam Hubungan Dengan Produksi Padi Sawah (Oryza sativa L.) Di Kabupaten Minahasa Utara," *Agri-Sosioekonomi*, vol. 19, no. 1, pp. 599–604, 2023, doi: 10.35791/agrsosek.v19i1.46854.
- [3] P. C. Latue, J. Steiven, I. Septory, G. Somae, and H. Rakuasa, "PEMODELAN DAERAH RAWAN BANJIR DI KECAMATAN SIRIMAU MENGGUNAKAN METODE MULTI-CRITERIA ANALYSIS (MCA) Modeling Flood-Prone Areas in Sirimau District Using the Multi-Criteria Analysis (MCA) Method," vol. 18, pp. 10–17, 2023.
- [4] I. Fatima and D. P. Sudiby, "Efektivitas Badan Penanggulangan Bencana Daerah (BPBD) dalam mitigasi bencana erupsi



- Gunung Merapi di Kabupaten Magelang,” J. Mhs. Wacana Publik, vol. 3, no. 1, pp. 136–150, 2023.
- [5] H. Sharfina, P. Y. Utami, and I. Fakhruzi, “Prediksi Bencana Banjir Menggunakan Algoritma Deep Learning H2O Berdasarkan Data Curah Hujan,” vol. 10, no. 4, 2023.
- [6] E. T. Ena Tasia and M. Afdal, “Perbandingan Algoritma K-Means Dan K-Medoids Untuk Clustering Daerah Rawan Banjir Di Kabupaten Rokan Hilir,” *Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 65–73, 2023, doi: 10.57152/ijirse.v3i1.523.
- [7] M. B. Arya Darmawan, F. Dewanta, and S. Astuti, “Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan Naïve Bayes untuk Prediksi Banjir di Desa Dayeuhkolot,” *TELKA - Telekomun. Elektron. Komputasi dan Kontrol*, vol. 9, no. 1, pp. 52–61, 2023, doi: 10.15575/telka.v9n1.52-61.
- [8] N. S. Fatonah, Penerapan Deteksi Bencana Banjir Menggunakan Metode Machine Learning, vol. 10, no. 2. 2021. doi: 10.22441/format.2021.v10.i2.002.
- [9] J. Nasir and P. S. Informasi, “PENERAPAN DATA MINING CLUSTERING DALAM MENGELOMPOKAN BUKU DENGAN METODE K-MEANS,” vol. 11, no. 2, 2020.
- [10] Y.-L. Cheng et al., We are IntechOpen , the world ’ s leading publisher of Open Access books Built by scientists , for scientists TOP 1 % , vol. 11, no. tourism. 2020. [Online]. Available: <https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics>
- [11] S. Nurajizah and A. Salbinda, “Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten,” vol. 7, no. 2, pp. 158–163, 2021, doi: 10.31294/jtk.v4i2.
- [12] C. Sukmayadi, A. Primajaya, and I. Maulana, “Penerapan Algoritma K-Medoids dalam Menentukan Daerah Rawan Banjir di Kabupaten Karawang,” vol. 6, no. 3, pp. 187–196, 2021.
- [13] S. Fatonah, “Penerapan Deteksi Bencana Banjir Menggunakan Metode Machine Learning,” vol. 10, pp. 119–126, 2021.
- [14] E. Tasia and M. Afdal, “Comparison Of K-Means And K-Medoid Algorithms For Clustering Of Flood-Prone Areas In Rokan Hilir District Perbandingan Algoritma K-Means Dan K-Medoids Untuk Clustering Daerah Rawan Banjir Di Kabupaten Rokan Hilir,” vol. 3, no. 1, pp. 65–73, 2023.
- [15] N. Septiani, “Penerapan Algoritma K-Means Clustering Untuk Harga Rumah di Jakarta Selatan,” vol. 1, no. 2, 2023.
- [16] D. Kurniadi, Y. H. Agustin, H. Ilham, N. Akbar, and I. Farida, “Penerapan Algoritma k-Means Clustering untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang,” vol. 20, no. 1, pp. 64–77, 2023.
- [17] W. W. Pribadi et al., “PERBANDINGAN METODE K-MEANS EUCLIDEAN DISTANCE DAN MANHATTAN DISTANCE PADA PENENTUAN ZONASI COVID-19 DI,” vol. 6, no. 2, 2022.
- [18] M. R. Nahjan, N. Heryana, A. Voutama, F. I. Komputer, U. S. Karawang, and R. Miner, “IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL,” vol. 7, no. 1, pp. 101–104, 2023.
- [19] A. Badruttamam and D. A. I. Maruddani, “PENERAPAN ANALISIS KLASTER K -MODES DENGAN VALIDASI DAVIES BOULDIN INDEX DALAM MENENTUKAN KARAKTERISTIK KANAL YOUTUBE DI INDONESIA (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade),” vol. 9, pp. 263–272, 2020.
- [20] A. Aranta, “ANALISIS PEMILIHAN CLUSTER OPTIMAL DALAM SEGMENTASI,” vol. 18, no. 2, pp. 152–163, 2021.