

KLASTERISASI WILAYAH DKI JAKARTA BERDASARKAN DATA KEJADIAN BENCANA PEMPROV DKI JAKARTA MENGUNAKAN K-MEANS

Widyo Isworo Nugroho¹, Yuliazmi^{2*}

¹Sistem Informasi, Fakultas Teknik Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹widyoisworo15@gmail.com, ^{2*}yuliazmi@budiluhur.ac.id,
(* : corresponding author)

Abstrak-Bencana adalah peristiwa atau rangkaian peristiwa yang dapat menimbulkan penderitaan manusia, kerusakan harta benda, kerusakan lingkungan, sarana dan prasarana, serta mempengaruhi kehidupan orang lain, kehidupan manusia dan penghidupan. Negara Kesatuan Republik Indonesia merupakan negara kesatuan dengan wilayah yang luas dan terletak di garis khatulistiwa dimana dua benua dan dua samudra bertemu, kondisi alam memiliki banyak keunggulan. Namun disisi lain, letaknya di wilayah geografis, geologis, hidrologis, dan demografis rawan bencana dan memerlukan pengelolaan yang sistematis, global, dan terkoordinasi dengan frekuensi tinggi. Pengelompokan Bencana pada suatu wilayah akan mempermudah pemerintah dalam pertimbangan di wilayah mana saja yang memerlukan atensi khusus dalam penanggulangan bencana alam di Indonesia. Dengan adanya pengelompokan tersebut akan dapat dilihat wilayah mana saja yang berdampak pada bencana alam. Data yang diperoleh dengan kurun waktu Januari 2021-juli 2021. Pada penelitian kali ini teknik dari data *mining* untuk klasterisasi wilayah di DKI Jakarta yaitu dengan menggunakan algoritma K-Means dan menggunakan metode CRISP-DM. Data dari penelitian ini didapat dan bersumber dari *website* data.jakarta.go.id. Dataset hanya data di provinsi DKI Jakarta. Yang terdiri dari tanggal, waktu kejadian, lokasi bencana, jenis bencana, jumlah korban bencana, persentase korban bencana yang mendapat bantuan sosial dan jenis bantuan yang diberikan. Berdasarkan hasil dari penelitian ini menunjukkan bahwa algoritma K-Means dapat menentukan 4 (empat) *cluster* yang optimal yaitu *cluster* satu dengan jumlah 31 *items*, *cluster* dua dengan jumlah 1 *items*, *cluster* tiga dengan jumlah 40 *items* dan *cluster* empat dengan jumlah 64 *items*. Dengan nilai Davies Bouldin (DB) algoritma K-Means sebesar 0,323 dibanding nilai DB algoritma X-Means sebesar 0,326. Hasil yang diperoleh dapat menjadi masukan bagi pemerintah di Provinsi DKI Jakarta dalam menganalisis jumlah kejadian bencana berdasarkan wilayah yang terdampak bencana.

Kata Kunci: Data Mining, K-Means, Bencana, Algoritma Clustering

CLUSTERIZATION OF THE DKI JAKARTA AREA BASED ON DISASTER INCIDENCE DATA OF THE PROVINCE OF DKI JAKARTA USING K-MEANS

Abstract- Disaster is an event or series of events that can cause human suffering, property damage, environmental damage, facilities and infrastructure, and affect the lives of other people, human life and livelihoods. The Unitary State of the Republic of Indonesia is a unitary state with a vast territory and is located on the equator where two continents and two oceans meet, natural conditions have many advantages. However, on the other hand, it is located in a geographical, geological, hydrological and demographic area which is prone to disasters and requires systematic, global and coordinated management with high frequency. Categorizing disasters in an area will make it easier for the government to consider which areas require special attention in managing natural disasters in Indonesia. With this grouping, it can be seen which areas have an impact on natural disasters. The data were obtained for the period January 2021-July 2021. In this research, the data mining technique for regional clustering in DKI Jakarta is by using the K-Means algorithm. The data from this study were obtained and sourced from the *website* data.jakarta.go.id. The dataset is only data in the province of DKI Jakarta. Which consists of the date, time of incident, location of the disaster, type of disaster, number of disaster victims, percentage of disaster victims who received social assistance and type of assistance provided. Based on the results of this study, it shows that the K-Means algorithm can determine the optimal 4 (four) clusters, namely cluster one with a total of 31 items, cluster two with a total of 1 item, cluster three with a total of 40 items and cluster four with a total of 64 items. With a Davies Bouldin (DB) value of the K-Means algorithm of 0.323 compared to the X-Means algorithm DB value of 0.326. The results obtained can be used as input for the government in DKI Jakarta Province in analyzing the number of disaster events based on the areas affected by the disaster.

Keywords: Data Mining, K-Means, Disaster, Clustering Algorithm

1. PENDAHULUAN

Negara Kesatuan Republik Indonesia adalah negara kesatuan dengan wilayah yang luas dan terletak di garis khatulistiwa, tempat bertemunya dua benua dan dua samudera, kondisi alam yang mendukung. Namun disisi lain, lokasi tersebut berada pada wilayah geografis, geologis, hidrologis, dan demografis yang sering terjadi bencana alam, sehingga memerlukan perhatian yang sistematis, komprehensif, dan terkoordinasi [2].

Bencana adalah peristiwa atau rangkaian peristiwa yang dapat menimbulkan penderitaan manusia, kerusakan harta benda, kerusakan lingkungan, sarana dan prasarana, serta mempengaruhi kehidupan orang lain, kehidupan manusia dan penghidupan. Lingkungan kita khususnya di Pulau Jawa semakin hari semakin rusak dan nampaknya sangat “sensitif” terhadap gangguan proses alam, seperti hujan lebat disertai angin kencang atau seringnya terjadi badai [1].

Penambangan data atau data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi yang berguna dan pengetahuan terkait dari database besar [3]. Penambangan data merupakan salah satu bagian dari kecerdasan buatan. Teknik dalam data mining yaitu asosiasi, clustering, estimasi, dan prediksi. Teknik penambangan data yang dipakai dalam penelitian ini yaitu clustering.

Clustering atau analisis cluster adalah proses pembentukan kelompok data (cluster) dari kumpulan data kelompok atau kelas yang tidak diketahui, dan proses penentuan data mana yang termasuk dalam cluster mana. Dari perspektif penambangan data, pengelompokan bukanlah proses klarifikasi. Karena pada saat klarifikasi, data dikelompokkan ke dalam kelas-kelas yang telah diketahui sebelumnya [4].

Beberapa penelitian terdahulu menunjukkan bahwa data mining dapat digunakan dengan baik dalam mengelompokkan data bencana alam[5], Penelitian ini mengklasifikasikan data bencana dengan teknik clustering menggunakan algoritma k-means ke dalam kelompok-kelompok menurut jenis bencana, waktu bencana, jumlah korban dan kerusakan fasilitas pasca bencana. Dataset ini digunakan untuk mengidentifikasi daerah rawan bencana, mengidentifikasi jenis bencana yang biasa terjadi di suatu daerah, mengidentifikasi kejadian bencana pada bulan tertentu, jumlah korban luka, jumlah korban meninggal, serta kerusakan yang dihasilkan. [6].

Terdapat metode-metode dari penelitian terdahulu yang sudah melakukan penentuan strategi promosi clustering terhadap jenis bencana dengan menggunakan metode algoritma K-Means [7].

Pada penelitian ini teknik dari data mining untuk mengelompokkan wilayah di DKI Jakarta yaitu dengan menggunakan algoritma K-Means. Algoritma K-Means clustering mempunyai kelebihan yaitu mudah dipahami dan diimplementasi untuk suatu penelitian. Dalam mempelajari algoritma tersebut membutuhkan waktu yang relatif cepat. Pada penelitian ini metode K-Means akan diterapkan untuk melakukan clustering wilayah di DKI Jakarta sehingga dapat mengetahui wilayah dengan jumlah bencana yang tinggi dan rendah. Penentuan jumlah cluster yang optimal menggunakan Elbow method.

Elbow Method adalah metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster yang akan membentuk siku pada suatu titik.

2. METODE PENELITIAN

Metode CRISP-DM

CRISP-DM (Cross Industry Standard Process Model for Data Mining) adalah metode yang banyak digunakan oleh para profesional menggunakan proses pemodelan data. Tujuan dari metode CRISP-DM adalah menemukan pola yang menarik dan penting dalam data yang digunakan. CRISP-DM memiliki langkah-langkah dan framework yang terstruktur sehingga pengguna metode lebih fokus dan mengetahui langkah-langkah yang harus dilakukan dalam penelitian [8]. Pada tahapan ini menerapkan metode CRISP-DM (Cross Industry Standard Process Model for Data Mining) dengan beberapa tahapan didalamnya. Model Proses CRISP-DM dapat dijelaskan sebagai berikut:



Gambar 1. Tahapan Metode CRISP-DM

Berdasarkan Gambar 1, berikut adalah penjelasan tentang tahapan-tahapan yang dilakukan :

- a. *Business Understanding*
Pada tahap awal, subjek penelitian ini adalah seluruh wilayah DKI Jakarta. Oleh karena itu, perlu adanya strategi penjumlahan atau agregat data untuk seluruh wilayah DKI Jakarta.
- b. *Data Understanding*
Peneliti melakukan survei kebutuhan data mengenai korelasi kejadian bencana di seluruh wilayah DKI Jakarta. Penelitian ini menggunakan data dinas sosial dari data.jakarta.go.id. Dataset hanya mencakup data provinsi dari DKI Jakarta. Ini meliputi tanggal dan waktu kejadian, lokasi bencana, jenis bencana, jumlah korban, persentase korban yang menerima bantuan sosial, dan jenis bantuan yang diberikan.
- c. *Data Preparation*
Persiapan data dilakukan dalam beberapa proses antara lain pemilihan variabel untuk analisis, pembersihan data, persiapan data awal hingga data siap untuk transformasi. Data seleksi diambil dari dinas sosial. Selanjutnya, melakukan preprocessing data. dalam proses ini bertujuan untuk melihat kualitas data yang dipilih dalam proses pemilihan data. Pada tahap ini dilakukan proses pembersihan data, yaitu proses pembersihan data dari nilai yang hilang atau noise. Langkah selanjutnya adalah melakukan transformasi data menggunakan metode normalisasi agar tidak terlalu banyak perbedaan kecepatan data.
- d. *Modeling*
Pada tahap pemodelan, pemilihan dan penerapan teknik pemodelan dan parameter tertentu akan disesuaikan untuk mendapatkan *cluster* yang optimal. Ada beberapa model untuk memecahkan masalah yang sama dalam data mining. Beberapa model ini memiliki spesifikasi dan format data tertentu. Pada penelitian ini akan digunakan algoritma K-Means. K-Means merupakan salah satu algoritma pelatihan unsupervised, awal mula algoritma ini dipublikasikan oleh Stuart Lloyd di tahun 1984 yang hingga saat ini merupakan algoritma yang banyak digunakan. Alasan algoritma ini banyak digunakan adalah karena K-Means cukup mudah untuk diimplementasikan, relatif cepat, dan mudah dimodifikasi [9].
- e. *Evaluation*
Pada tahap ini penekanan pada model yang dihasilkan konsisten dengan kriteria K-Means Clustering, dan tidak ada yang tertinggal pada tahap awal hingga tahap pemodelan selesai. Metode penilaian hasil secara cluster dalam metode clustering menggunakan metode indeks Davies-Bouldin. Dasar dari pendekatan ini adalah nilai koherensi dan pemisahan. Selama pengelompokan, koherensi adalah jumlah jarak antara data dan pusat cluster yang dipantau. Sedangkan pemisahan didasarkan pada jarak antar pusat cluster [10].
- f. *Deployment*
Pada tahap akhir ini, CRISP-DM merupakan diseminasi pengetahuan atau informasi yang diperoleh, diimplementasikan dalam bentuk laporan dan analisis dari setiap klaster yang diperoleh sehingga dapat dengan mudah dipahami oleh pihak yang membutuhkan.

3. HASIL DAN PEMBAHASAN

3.1 Business Understanding

Penelitian ini menggunakan data dari Dinas Sosial DKI Jakarta yang diperoleh dari website <https://data.jakarta.go.id/dataset/> pada kurun waktu semester genap 2021. Data tersebut terdiri dari wilayah-wilayah yang terkena bencana di DKI Jakarta meliputi data tanggal mulai, tanggal selesai, jam kejadian, lokasi bencana, kelurahan, kecamatan, wilayah, jenis bencana, jumlah korban bencana, persentase korban bencana yang mendapat bantuan, dan jenis bantuan yang diberikan.

3.2 Data Preparation

Pada tahapan ini data yang didapat sebanyak 136 record kejadian bencana pada tahun 2021. Tahapan ini yang diterapkan pada database untuk menghapus noise, cleaning data, missing value dan data yang tidak konsisten. Data preprocessing pada penelitian ini penulis melakukan data transformation. Dalam data realtime database terkadang masih belum lengkap. Langkah-langkah yang terdapat dalam tahap ini tidak semuanya harus dilakukan. Berikut langkah-langkah yang dilakukan penulis :

1. *Data Transformation*

Pada tahapan ini untuk mengubah seluruh data yang terkumpul agar memudahkan ketika diolah dengan menggunakan metode K-Means clustering. Terdapat beberapa teknik untuk melakukan data transformation adalah normalization, pemilihan attribute, dan discretization. Adapun variabel yang

digunakan pada penelitian ini yaitu Jam Kejadian, Kelurahan, Kecamatan, Wilayah, Jenis Bencana dan Jumlah Korban.

Berikut beberapa atribut yang ditransformasikan dengan label encoding :

1. Jam Kejadian

Tabel 1 Hasil Transformasi Data Jam Kejadian

Jam Kejadian	Kode
Dini hari (00:01 – 06:00)	1
Pagi hari (06:01 – 12:00)	2
Siang hari (12:01 – 15:00)	3
Sore hari (15:01-18:00)	4
Malam hari (18:01 – 24:00)	5

2. Wilayah

Tabel 2 Hasil Transformasi Data Wilayah

Wilayah	Kode
Jakarta Barat	31.73
Jakarta Pusat	31.71
Jakarta Selatan	31.74
Jakarta Timur	31.75
Jakarta Utara	31.72

3. Wilayah Kecamatan Jakarta Barat

Tabel 3 Hasil Transformasi Data Wilayah Kecamatan Jakarta Barat

Kecamatan	Kode
Cengkareng	31.73.01
Kalideres	31.73.06
Kebon Jeruk	31.73.05
Kembangan	31.73.08
Palmerah	31.73.07
Taman Sari	31.73.03
Tambora	31.73.04

4. Wilayah Kecamatan Jakarta Pusat

Tabel 4 Hasil Transformasi Data Wilayah Kecamatan Jakarta Pusat

Kecamatan	Kode
Gambir	31.71.01
Kemayoran	31.71.03
Menteng	31.71.06
Sawah Besar	31.71.02
Senen	31.71.04
Tanah Abang	31.71.07

5. Wilayah Kecamatan Jakarta Selatan

Tabel 5 Hasil Transformasi Data Wilayah Kecamatan Jakarta Selatan

Kecamatan	Kode
Cilandak	31.74.06
Jagakarsa	31.74.09
Kebayoran Baru	31.74.07
Kebayoran Lama	31.74.05
Mampang Prapatan	31.74.03
Pancoran	31.74.08
Pasar Minggu	31.74.04
Pesanggrahan	31.74.10
Setiabudi	31.74.02
Tebet	31.74.01

6. Wilayah Kecamatan Jakarta Timur

Tabel 6 Hasil Transformasi Data Wilayah Kecamatan Jakarta Timur

Kecamatan	Kode
Cakung	31.75.06
Cipayung	31.75.10
Ciracas	31.75.09
Duren Sawit	31.75.07
Jatinegara	31.75.03
Kramat Jati	31.75.04
Makasar	31.75.08
Matraman	31.75.01
Pasar Rebo	31.75.05

7. Wilayah Kecamatan Jakarta Utara

Tabel 7 Hasil Transformasi Data Wilayah Kecamatan Jakarta Utara

Kecamatan	Kode
Cilincing	31.72.04
Kelapa Gading	31.72.06
Koja	31.72.03
Pademangan	31.72.05
Penjaringan	31.72.01
Tanjung Priok	31.72.02

8. Wilayah Kelurahan Jakarta Barat

Tabel 8 Hasil Transformasi Wilayah Kelurahan Jakarta Barat

Kelurahan	Kode
Cengkareng Barat	31.73.01.1001
Duri Kosambi	31.73.01.1002
Kedaung Kaliangke	31.73.01.1004
Rawa buaya	31.73.01.1003
Kamal	31.73.06.1004
Tegal Alur	31.73.06.1003
Kebon Jeruk	31.73.05.1001
Kedoya Selatan	31.73.05.1007
Kedoya Utara	31.73.05.1006
Joglo	31.73.08.1005
Kembangan Selatan	31.73.08.1006
Kembangan Utara	31.73.08.1001
Meruya Selatan	31.73.08.1003
Srengseng	31.73.08.1004
Kota Bambu Selatan	31.73.07.1006
Palmerah	31.73.07.1001
Keagungan	31.73.03.1006
Angke	31.73.04.1007
Tanah Sereal	31.73.04.1004

9. Wilayah Kelurahan Jakarta Pusat

Tabel 9 Hasil Transformasi Wilayah Kelurahan Jakarta Pusat

Kelurahan	Kode
Cideng	31.71.01.1002
Petojo Selatan	31.71.01.1004
Gunung Sahari Selatan	31.71.03.1005
Menteng	31.71.06.1001
Mangga Dua Selatan	31.71.02.1005
Kwitang	31.71.04.1005
Bendungan Hilir	31.71.07.1002
Karet Tengsin	31.71.07.1003
Kebon Melati	31.71.07.1005
Kebon Kosong	31.71.03.1002

10. Wilayah Kelurahan Jakarta Selatan

Tabel 10 Hasil Transformasi Wilayah Kelurahan Jakarta Selatan

Kelurahan	Kode
Gandaria Selatan	31.74.06.1004

Pondok Labu	31.74.06.1003
Jagakarsa	31.74.09.1001
Gunung	31.74.07.1002
Selong	31.74.07.1004
Rawa barat	31.74.07.1005
Cipete Utara	31.74.07.1010
Cipete Selatan	31.74.06.1005
Petogogan	31.74.07.1008
Kramat Pela	31.74.07.1003
Gandaria Utara	31.74.07.1009
Cipulir	31.74.05.1003
Kuningan Barat	31.74.03.1005
Tegal Parang	31.74.03.1005
Rawajati	31.74.08.1003
Jati padang	31.74.04.1002
Pasar Minggu	31.74.04.1001
Ragunan	31.74.04.1004
Pejaten Barat	31.74.04.1006
Kebagusan	31.74.04.1007
Bintaro	31.74.10.1002
Kalibata	31.74.08.1002
Karet Semanggi	31.74.02.1002
Menteng Dalam	31.74.01.1003
Kebon Baru	31.74.01.1004
Manggarai	31.74.01.1007

11. Wilayah Kelurahan Jakarta Timur

Tabel 11 Hasil Transformasi Wilayah Kelurahan Jakarta Timur

Kelurahan	Kode
Rawa terate	31.75.06.1002
Jatinegara	31.75.06.1001
Cakung Timur	31.75.06.1004
Lubang Buaya	31.75.10.1007
Setu	31.75.10.1005
Bambu Apus	31.75.10.1006
Cibubur	31.75.09.1002
Kelapa Dua Wetan	31.75.09.1003
Ciracas	31.75.09.1001
Susukan	31.75.09.1004
Kampung Rambutan	31.75.09.1005
Pondok Bambu	31.75.07.1002
Pondok Kelapa	31.75.07.1004
Bidara Cina	31.75.03.1002
Cipinang Besar Selatan	31.75.03.1007
Kampung Melayu	31.75.03.1001
Cipinang Muara	31.75.03.1006
Rawa bunga	31.75.03.1004
Cawang	31.75.04.1007
Cililitan	31.75.04.1006
Kramat Jati	31.75.04.1001
Batu Ampar	31.75.04.1004
Dukuh	31.75.04.1003
Balekambang	31.75.04.1005
Makasar	31.75.08.1001
Cipinang Melayu	31.75.08.1005
Pinang Ranti	31.75.08.1002
Kebon Pala	31.75.08.1003

Halim Perdanakusuma	31.75.08.1004
Kebon Manggis	31.75.01.1005
Palmeriam	31.75.01.1004
Pisangan Baru	31.75.01.1001
Baru	31.75.05.1002
Pekayon	31.75.05.1005
Cijantung	31.75.05.1003
Ujung Menteng	31.75.06.1006

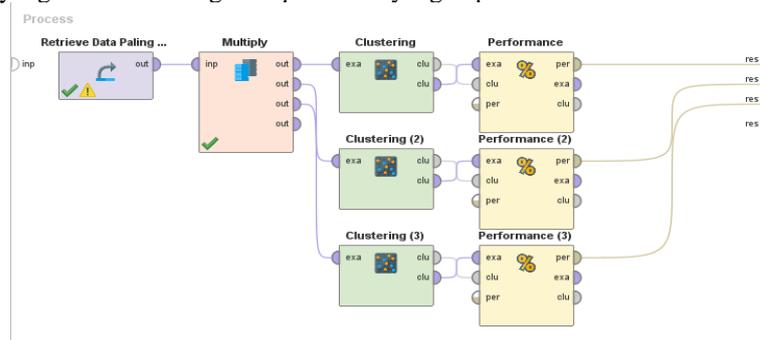
12. Wilayah Kelurahan Jakarta Utara

Tabel 12 Hasil Transformasi Wilayah Kelurahan Jakarta Utara

Kelurahan	Kode
Rorotan	31.72.04.1006
Semper Timur	31.72.04.1005
Kelapa Gading Barat	31.72.06.1003
Tugu Selatan	31.72.03.1005
Pademangan Timur	31.72.05.1001
Pejagalan	31.72.01.1004
Tanjung Priok	31.72.02.1001
Kebon Bawang	31.72.02.1005

3.3 Modeling

Pada tahap ini peneliti melakukan tahap modeling untuk menentukan cluster yang optimal menggunakan *k-means*. Berikut proses yang dilakukan dengan *RapidMiner* yang dapat dicermati di Gambar 2.



Gambar 2 Proses dari Algoritma K-Means

Setelah melakukan proses rapidminer, maka dapat dilihat hasil pada tabel 13 nilai DBI paling kecil dengan penentuan jumlah cluster terbaik adalah berjumlah 4 cluster dengan nilai 0,323. Semakin kecil nilai DBI maka semakin optimal cluster yang dihasilkan.

Tabel 13 Hasil Davies Bouldin Index

Penentuan Jumlah Cluster Terbaik	Hasil Davies Bouldin Index
2 Cluster	0,425
3 Cluster	0,479
4 Cluster	0,323

Selanjutnya, gambar 3. merupakan hasil 4 cluster yang diperoleh :

Cluster Model

```
Cluster 0: 31 items
Cluster 1: 1 items
Cluster 2: 40 items
Cluster 3: 64 items
Total number of items: 136
```

Gambar 3 Hasil Jumlah Cluster

Clustering seperti pada gambar 4.2 diatas, dengan jumlah data 136 dan inialisasi jumlah cluster sebanyak 4 buah, sesuai dengan pendefinisian nilai k dengan jumlah *cluster_0* : 31 items, *cluster_1* : 1 items, *cluster_2* : 40 items dan *cluster_3* : 64 items.

Dari data hasil penelitian clustering yang telah dilakukan diatas. Dapat disimpulkan bahwa *cluster_0* bencana yang sering terjadi adalah Banjir dan Kebakaran tetapi untuk bencana yang dominan ialah Banjir, dengan dominan wilayah di Jakarta Selatan, dan waktu kejadian dominan pada Sore Hari, dengan jumlah korban minimal 3 jiwa, maksimal 1116 jiwa serta rata-rata korban ialah 368 jiwa.

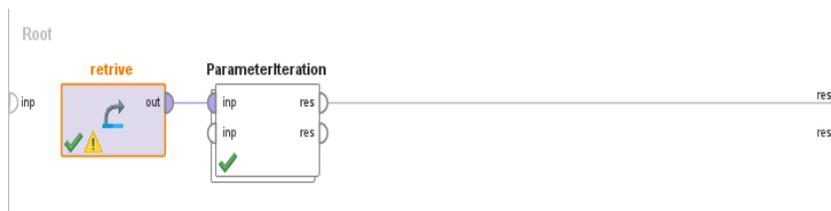
Pada *cluster_1* bencana yang dominan terjadi adalah Banjir, dengan wilayah di Jakarta Selatan dan waktu kejadian pada Sore Hari, dengan jumlah korban mencapai 5862 jiwa.

Selanjutnya *cluster_2* bencana yang dominan ialah Banjir dan Kebakaran, dengan dominan wilayah di Jakarta Barat dan waktu kejadian dominan pada Sore Hari, dengan jumlah korban minimal 4 jiwa, maksimal 1500 jiwa serta rata-rata korban ialah 277 jiwa.

Terakhir pada *cluster_3* bencana yang sering terjadi adalah Banjir, Kebakaran, dan Angin Puting Beliung tetapi untuk bencana yang dominan ialah Banjir, dengan dominan wilayah di Jakarta Timur, dan waktu kejadian dominan pada Pagi Hari, dengan jumlah korban minimal 6 jiwa, maksimal 1200 jiwa serta rata-rata korban ialah 281 jiwa.

3.4 Evaluation

Pada bagian ini disampaikan mekanisme pengukuran performa model terbaik. Metode uji dari penelitian ini menggunakan *Davies-Bouldin Index*. Berikut proses dari pengujian yang menggunakan *tools Rapidminer* :



Gambar 5 Operator Pengujian

Proses diatas menggunakan operator *loop parameters*. Lali dari hasil *processLog* menunjukkan nilai DB paling optimal ialah 0,323 dengan jumlah 4 cluster.

Berikut gambar 4.4 yang menunjukkan hasil pengujian dengan nilai DB paling optimal :

ProcessLog (11 rows, 3 columns)

k	DB	W
2	0.425	1249834...
3	0.479	653401....
4	0.323	438369....

Gambar 6 Hasil Pengujian

4. KESIMPULAN

1. Berdasarkan hasil penelitian yang ada dapat disimpulkan bahwa terdapat 4 *cluster*.
2. Berdasarkan hasil perhitungan nilai *Davies Bouldin Index* menggunakan *tools Rapidminer* dengan algoritma *K-Means* paling kecil dalam penentuan *cluster* terbaik adalah berjumlah 4 *cluster* yaitu 0,323.
3. Pada rentang tahun 2021 pada *cluster 0* di Jakarta Selatan memiliki kejadian bencana alam tertinggi pada kecamatan Kebayoran Lama dengan kejadian bencana alam dominan Banjir dan korban bencana di *Cluster 0* sebanyak 11401 jiwa. *Cluster 1* di Jakarta Selatan memiliki kejadian bencana alam tertinggi pada kecamatan Pasar Minggu dengan kejadian bencana alam dominan Banjir dan korban bencana di *Cluster 1* sebanyak 5862 jiwa. *Cluster 2* di Jakarta Barat dan Jakarta Pusat memiliki kejadian bencana alam tertinggi, tetapi dominan di wilayah Jakarta Barat dengan kejadian bencana tertinggi di kecamatan Cengkareng dengan kejadian bencana alam dominan Banjir dan korban bencana di *Cluster 2* sebanyak 11072 jiwa. *Cluster 3* di Jakarta Timur dan Jakarta Utara memiliki kejadian bencana alam tertinggi, tetapi dominan di wilayah Jakarta Timur dengan kejadian bencana tertinggi di kecamatan Kramat Jati dengan kejadian bencana alam dominan Banjir dan korban bencana di *Cluster 3* sebanyak 17987 jiwa.
4. Berdasarkan hasil penelitian yang didapat kekurangan dari penelitian ini adalah wilayah yang diteliti hanya mencakup provinsi DKI Jakarta saja dan metode yang digunakan hanya menggunakan *K-Means Clustering*.

UCAPAN TERIMA KASIH

Terima kasih kepada pihak-pihak yang sudah membantu yaitu orang tua dan keluarga saya serta teman-teman seperjuangan yang sudah membantu dalam memberikan inspirasi dalam pembuatan jurnal ini.

DAFTAR PUSTAKA

- [1] F. Umeidini, E. Nuriah, and M. Fedryansyah, "Di Desa Mekargalih Kecamatan Jatinangor," vol. 2, pp. 13–22, 2019.
- [2] A. Chandra, P. S. Geografi, F. Geografi, and U. M. Surakarta, "Analisis Kesiapsiagaan-Masyarakat-terhadap-Bencana Tanah- Longsor di Kecamatan-Tawangmangu-Kabupaten Karanganyar," 2023.
- [3] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [4] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.
- [5] Isni Rinjani, Saeful Anwar, and Ruli Herdiana, "PENGELOMPOKAN DAERAH BENCANA ALAM MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," *J. Ilm. Sist. Inf. dan Ilmu Komput.*, vol. 3, no. 1 SE-Articles, pp. 35–51, Mar. 2023, doi: 10.55606/juisik.v3i1.417.
- [6] M. Murdiaty, A. Angela, and C. Sylvia, "Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 744, 2020, doi: 10.30865/mib.v4i3.2213.
- [7] Y. Pratama, E. Rasywir, B. Tio Carenina, and D. Riski Angraini, "Penerapan Algoritma K-Means clustering Untuk Mengelompokkan Provinsi Berdasarkan Banyaknya Desa/Kelurahan Dengan Upaya Antisipasi/Mitigasi Bencana Alam," *Technol. Sci.*, vol. 4, no. 3, pp. 1232–1240, 2022, doi: 10.47065/bits.v4i3.2549.
- [8] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, "Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM," *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [9] A. A. Hussein, "Improve The Performance of K-means by using Genetic Algorithm for Classification Heart Attack," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 2, p. 1256, 2018, doi: 10.11591/ijece.v8i2.pp1256-1261.
- [10] R. K. Dinata, H. Novriando, N. Hasdyna, and S. Retno, "Reduksi Atribut Menggunakan Information Gain," vol. 6, no. 1, pp. 48–53, 2020.