

KLASTERISASI DATA BENCANA ALAM DI KOTA CIREBON MENGGUNAKAN K-MEANS CLUSTERING BERDASARKAN KAWASAN DAN JENIS BENCANA

Rahadatul Aisy¹, Rudi Kurniawan²

¹Manajemen Informatika, STMIK IKMI Cirebon

²Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jalan Perjuangan No. 10B Karyamulya Kec. Kesambi Kota Cirebon

aisysisy21@gmail.com

ABSTRAK

Bencana merupakan peristiwa yang memiliki dampak kerugian yang dirasakan oleh masyarakat seperti hancurnya rumah, bangunan, rusaknya fasilitas jalan, serta hilangnya pekerjaan. Bencana tersebut disebabkan oleh banyak faktor salah satunya adalah faktor dari manusia dan alam. Faktor yang disebabkan oleh non alam yaitu, kebakaran rumah, pohon tumbang, rumah ambruk dan untuk faktor bencana alamnya adalah tanah longsor, gempa bumi, dan banjir. Untuk memiliki sikap sigap terhadap bencana yaitu bisa dengan cara identifikasi kawasan atau wilayah yang rentan terjadinya bencana, khususnya pada wilayah kota Cirebon. Saat ini BPBD belum memiliki pengelompokan data terhadap Kawasan yang rentan bencana dengan jenis bencana yang telah terjadi di kota Cirebon sehingga kurang efektif dalam mengetahui kawasan dan jenis bencana yang rentan terjadi. Permasalahan ini memuat pengelompokan kawasan rentan bencana dan jenis bencana di kota Cirebon menggunakan metode *K-means*. Metode *K-means clustering* adalah untuk mengkategorikan pengelompokan sekelompok objek sesuai dengan atribut yang sama atau karakteristik ke dalam sejumlah grup. Penelitian ini menerapkan *K-means clustering* tersebut untuk mengelompokkan kawasan kejadian bencana alam berdasarkan jenis bencana di kota Cirebon. Tujuan klasterisasi ini untuk mengetahui kawasan-kawasan di kota Cirebon yang rentan terjadi bencana berdasarkan jenis bencana yang ada. Hasil dari klasterisasi menghasilkan 6 kelompok dengan anggota yang berbeda dan nilai DBI yang diperoleh 0.087.

Kata kunci : Bencana alam, Jenis bencana, Klasterisasi, *K-means Clustering*, Kota Cirebon

1. PENDAHULUAN

Bencana alam adalah bencana yang disebabkan oleh faktor alam, non alam, dan manusia. Bencana alam adalah suatu kejadian atau kerusakan yang disebabkan oleh alam. Kejadian ini biasanya terjadi secara tiba-tiba atau mendadak dan berpotensi menyebabkan kerugian materi bahkan jiwa manusia. Salah satu fenomena alam yang mengancam keberlangsungan hidup manusia adalah bencana alam. Kerugian dapat berupa kerugian material atau nonmaterial. Bencana alam seperti banjir, tanah longsor, dan gempa bumi adalah contohnya, tetapi ada juga bencana non alam seperti kebakaran akibat kegagalan teknologi, ketinggalan zaman, konflik sosial, dan teror. Bencana adalah kejadian dalam kehidupan manusia yang tidak diketahui kapan terjadi. Orang-orang hanya dapat mengidentifikasi gejala awal dan memperkirakan perkembangan gejala berikutnya [1]

Adanya peristiwa bencana alam tentu mengakibatkan banyak kerugian bagi masyarakat. Dampak dari bencana alam yaitu, rusaknya infrastruktur, adanya penyakit, terganggunya aktivitas sehari-hari dan sulit untuk mengakses jalan. Banyak faktor yang mempengaruhi terjadinya bencana alam sehingga masyarakat harus menjaga lingkungan sekitar untuk mengurangi terjadinya bencana alam yang diakibatkan ulah manusia sendiri. Kota Cirebon menjadi salah satu kota yang sering terjadi bencana, hal ini dilihat dari informasi yang tercatat oleh Badan Penanggulangan Bencana Daerah (BPBD) Kota

Cirebon. BPBD adalah Lembaga pemerintah yang bertugas untuk menanggulangi bencana khususnya lingkup daerah. Salah satu tujuan dari BPBD adalah untuk mengedukasi masyarakat terhadap cara penanganan ketika terjadi bencana alam maupun non alam. Peristiwa bencana alam tidak bisa diprediksi kapan akan terjadi. Sehingga, masyarakat harus memiliki sikap siaga terhadap bencana. Tindakan tersebut bisa dilakukan dengan melihat data bencana terdahulu. Pada data bencana yang ada belum adanya pengelompokan kawasan rentan bencana dan jenis bencana yang terjadi di kota Cirebon. Sehingga kurang efektif dalam pengambilan keputusan manajemen bencana. Oleh karena itu, penelitian ini akan mengelompokkan data bencana di kota Cirebon dengan tujuan menghasilkan data di kawasan mana saja yang rentan terjadi bencana di kota Cirebon.

2. TINJAUAN PUSTAKA

2.1. Landasan Teori

Penelitian yang dilakukan oleh Khomsiyah tahun 2021 dengan judul Penerapan Algoritma *K-Means Clustering* Untuk Pengelompokan Wilayah Rawan Banjir pada tahun 2018 untuk menentukan wilayah rawan banjir. Metode yang digunakan Algoritma *K-Means Clustering*. Hasil penelitian menghasilkan n 5 cluster yang terdiri dari Cluster 0 berjumlah 8 items, Cluster 1 berjumlah 15 items, Cluster 2 berjumlah 33 items Cluster 3 berjumlah 10 items, dan Cluster 4 berjumlah 32 items.

2.2. K-Means

K-Means merupakan teknik mengelompokkan data non-hierarki (partisi) yang memiliki kemampuan untuk membagi data menjadi dua kelompok atau lebih. Data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam kelompok yang sama, dan data dengan karakteristik yang berbeda akan dikelompokkan ke dalam kelompok yang berbeda. Meminimalkan perbedaan dalam kelompok dan meminimalkan perbedaan antar kelompok adalah tujuan dari pengelompokan untuk mengurangi fungsi objektif yang ditetapkan sepanjang proses [4]

Pada penjelasan langkah Algoritma *K-Means* dapat dilihat melalui Gambar 2.2 yang diperoleh dari jurnal Analisis *Clustering* Kasus Covid 19 di Indonesia Menggunakan Algoritma *K-Means* oleh Tikaridha Hardiani [5]



Gambar 1. Flowchart Algoritma *K-Means*

Pada Gambar 1 merupakan langkah-langkah Algoritma *K-means*.

2.3. Clustering

Clustering adalah metode pengelompokan titik data ke dalam dua kelompok atau lebih sehingga titik data dalam kelompok yang sama lebih mirip satu sama lain daripada di kelompok yang berbeda berdasarkan informasi yang tersedia untuk titik data dalam kelompok tersebut. Metode *clustering* telah menjadi alat yang berguna untuk memecahkan masalah kompleks ilmu komputer dan statistic [6]

Salah satu metode data *mining* yang bersifat tanpa arahan adalah *clustering*. Dalam proses pengelompokan data, dua jenis data *clustering* yang paling umum adalah *hierarchical* (hirarki) dan *non-hierarchical* (non-hirarki). Identifikasi kelas objek yang memiliki kemiripan juga disebut sebagai *clustering*. Dengan menggunakan teknik *clustering*, dapat menemukan kepadatan dan jarak daerah objek ruang serta pola distribusi dan korelasi antara atribut [7]

Dalam *Clustering* terdapat dua jenis data pengelompokan:

- a. *Hierarchical Clustering*
Teknik *clustering* data *hierarchical* dikelompokkan dengan cara *hierarchical clustering*, yang melibatkan pembagian seluruh set data ke dalam grup atau penggabungan dua grup yang lebih dekat di setiap iterasi.
- b. *Non-Hierarchical Clustering*
Teknik *Clustering* data *non-hierarchical* menggabungkan data menjadi satu atau lebih klaster atau grup. *Cluster* atau grup ini menggabungkan data dengan atribut yang sama dan data dengan atribut yang berbeda, sehingga variabilitas data di dalamnya berkurang [8].

2.5. Rapidminer

Rapidminer adalah *platform* perangkat lunak ilmu data dan pembelajaran mesin yang kuat yang menawarkan berbagai alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. Desainnya yang mudah digunakan memungkinkan pengguna tanpa pengalaman pemrograman membangun dan menguji berbagai model. *Rapidminer* mendukung berbagai sumber data, termasuk *file* datar, basis data, dan platform *big data*. Perangkat lunak ini juga mencakup berbagai operator yang sudah dibangun, yang merupakan blok bangunan alur kerja yang mencakup semua tahap proses *data mining* [9].

3. METODE PENELITIAN

3.1. Teknik Pengumpulan Data

3.1.1. Studi Dokumentasi

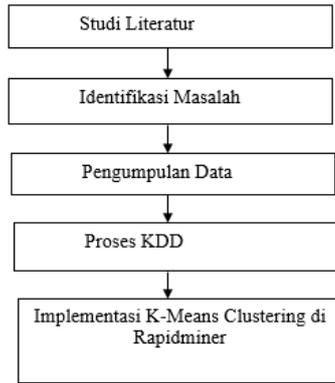
Metode studi dokumentasi digunakan untuk mendapatkan data bencana alam kota Cirebon yang diperoleh dari portal Cirebon Satu Data.

3.1.2. Studi Literatur

Teori studi literatur digunakan dalam penelitian Tugas Akhir ini untuk menyelesaikan masalah dan menjadi sumber referensi yang kuat. Peneliti menggunakan jurnal nasional dan internasional yang berkaitan dengan topik penelitian.

3.2. Tahapan Penelitian

Dalam penelitian, tahapan penelitian menjelaskan proses yang diambil mulai dari menemukan masalah hingga menemukan solusinya [10]. Gambar 3.1 berikut menunjukkan tahapan penelitian yang dilakukan



Gambar 2. Tahapan Penelitian

Gambar 2. adalah tahap proses penelitian pada tugas akhir.

3.3. Studi Literatur

Dalam tahap ini mempelajari berbagai referensi pada penelitian terdahulu yang relevan dengan penelitian yang Peneliti teliti mengenai Algoritma *K-Means Clustering*.

3.4. Identifikasi Masalah

Setelah mempelajari referensi terdahulu, kemudian memahami masalah serta mencari solusi dan metode untuk penelitian. Peneliti fokus terhadap permasalahan yang dibahas dalam latar belakang yaitu, kawasan rawan bencana di kota Cirebon masih belum adanya pengelompokan kawasan mana saja yang rentan sehingga sulit mengimplementasikan strategi mitigasi yang efektif. Diperlukan tindakan konkret untuk merancang dan menerapkan pengelompokan kawasan tersebut agar dapat lebih proaktif dalam menghadapi potensi risiko bencana.

3.5. Pengumpulan Data

Pada tahap ketiga dilakukan pengumpulan data yang diperoleh dari portal <https://data.cirebonkota.go.id/> milik pemerintah kota Cirebon yang menyediakan berbagai data akurat dari seluruh organisasi daerah kota Cirebon. Serta diperoleh dari pihak BPBD kota Cirebon. Data tersebut adalah data bencana alam kota Cirebon yang menunjukkan kejadian bencana di beberapa kawasan kota Cirebon berdasarkan jenis bencana nya pada periode 2019-2023 yang disimpulkan oleh Badan Penanggulangan Bencana Daerah (BPBD) kota Cirebon.

3.6. Proses KDD

Dalam tahap keempat perancangan ini adalah dilakukannya proses *Knowledge Discovery in Database* (KDD). KDD adalah proses yang diperlukan dalam data mining. Ini terdiri dari

langkah-langkah seperti pembersihan data, integrasi data, pemilihan data, transformasi data, penambahan data, evaluasi pola, dan presentasi pengetahuan.

3.7. Implementasi K-Means Clustering di Rapidminer

Pada tahap ini, pengujian data akan dilakukan dengan metode yang telah ditentukan. Sehingga, akan menentukan apakah hasil yang dihasilkan akan menjawab rumusan masalah atau sebaliknya.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Implementasi K-Means Menggunakan Rapidminer

Setelah melalui tahapan proses, kemudian data diolah menggunakan aplikasi Rapidminer dengan metode K-Means Clustering. Pengolahan data terdiri dari proses:

a. Data Selection

Dalam perancangan tugas akhir ini, dataset yang digunakan terdiri dari data bencana alam yang berisi kawasan yang rawan bencana alam berdasarkan jenis bencana yang terjadi pada periode Januari 2019-Desember 2023.

Tabel 1. Data Bencana Alam

Nama desa keluarahan	Jenis Bencana alam	Jumlah Kawasan	Tahun
ARGASUNYA	BANJIR	0	2023
ARGASUNYA	BANJIR ROB	0	2023
ARGASUNYA	LONGSOR	0	2023
ARGASUNYA	GEMPA BUMI	1	2023
ARGASUNYA	CUACA EKSTREM	1	2023
...
KESENDEN	LETUSAN GUNUNG MERAPI	0	2019

b. Data Cleaning

Proses pembersihan data adalah proses untuk mengisi nilai yang tidak ada atau tidak sesuai, mengoreksi data yang tidak konsisten, dan mendeteksi redudansi data, yaitu penumpukan atau duplikasi data.

c. Data Transformation

Proses mengubah data untuk mendapatkan data yang lebih besar atau sesuai dengan kebutuhan pengguna dikenal sebagai transformasi data. Dalam algoritma *K-Means Clustering*, data yang dapat diproses adalah data numerik. Berikut Tabel 2. adalah data hasil Transformation.

Tabel 2. Hasil Data Transformation

Kelurahan	Banjir	Bannjir Rob	Longsor	Gempa Bumi	Cuaca Ekstrem	Kekeringan	Letusan Gunung Merapi	Tahun
Argasunya	0	0	0	1	0	1	0	2023
Kalijaga	2	0	0	1	0	0	0	2023
Harjamukti	1	0	2	1	1	0	0	2023
Kecapi	1	0	0	1	1	0	0	2023
....
Kesenden	1	0	0	0	0	0	0	2019

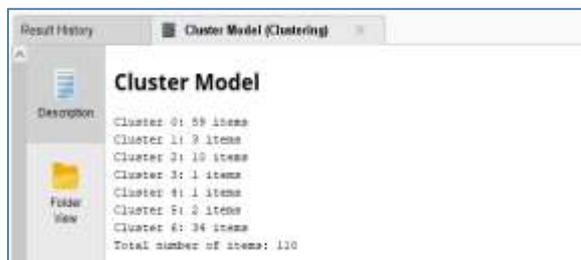
d. Data Mining

Pada proses Data Mining pertama yaitu proses mengelompokkan data berdasarkan jenis bencana dan kawasan. Dalam proses data mining digunakannya operator K-Means, sebelumnya memproses operator K-Means tentukan terlebih dahulu K berapa yang ingin digunakan. Pada proses ini menggunakan K=7 dan max run 10. Gambar 3. berikut menunjukkan tahap memasukan operator K-Means.



Gambar 3. Proses data mining

Setelah operator *K-Means* dijalankan, menghasilkan 7 cluster dari 110 dataset. Hasil *Clustering K-Means* dapat dilihat pada Gambar 4. Berikut.



Gambar 4. Cluster Model

Berdasarkan hasil pengelompokan data bencana alam menggunakan perhitungan *K-Means* dengan Rapidminer, dapat dipaparkan sebagai berikut:

- a. Cluster 6, merupakan cluster dengan frekuensi bencana alam sangat tinggi. Cluster 6 terdiri dari 34 anggota yaitu, Argasunya-Kesenden tahun 2023. Argasunya, Harjamukti, Larangan, Lemahwungkuk, Panjunan, Jagasatru, Pulasaren,

Pekalangan, Karyamulya, Sunyaragi, Kesambi, Kebon baru, Kesenden tahun 2022.

- b. Cluster 2, merupakan cluster dengan frekuensi bencana alam tinggi 1, terdiri dari 10 anggota yaitu, Sukapura tahun 2023, Kalijaga, Kecapi, Pekalipan, Pekiringan, Kejaksan, Sukapura tahun 2022, Sukapura tahun 2021 dan Kesepuhan tahun 2020.
- c. Cluster 3, merupakan cluster dengan frekuensi bencana alam sedang 2, terdiri dari 1 anggota yaitu, Kesepuhan tahun 2022.
- d. Cluster 1, merupakan cluster dengan frekuensi bencana alam sedang 1, yang terdiri dari 3 anggota yaitu, Drajat tahun 2022, Kalijaga tahun 2020 dan Pegambiran tahun 2020.
- e. Cluster 5, merupakan cluster dengan frekuensi bencana alam rendah, yang terdiri dari 2 anggota yaitu, Argasunya tahun 20220 dan Argasunya tahun 2021.
- f. Cluster 0, merupakan cluster dengan frekuensi bencana alam Tinggi 2, yaitu hanya terdiri 59 anggota yaitu, Kalijaga, Harjamukti, Kecapi, Larangan, Pegambiran, Kesepuhan, Panjunan, Jagasatru, Pulasaren, Pekalipan, Pekalangan, Karyamulya, Sunyaragi, Drajat, Kesambi, Pekiringan, Kejaksan, Kebon baru, Kesenden tahun 2021. Harjamukti, Kecapi, Larangan, Lemahwungkuk, Panjunan, Jagasatru, Pulasaren, Pekalipan, Pekalangan, Karyamulya, Sunyaragi, Drajat, Kesambi, Pekiringan, Kejaksan, Kebon baru, Sukapura, Kesenden tahun 2020. Argasunya-Kesenden tahun 2019.
- g. Cluster 4, merupakan cluster dengan frekuensi bencana alam sangat rendah, yaitu terdiri 1 anggota yaitu Lemahwungkuk tahun 2021.

4.2. Hasil Pengujian nilai K ideal dengan nilai DBI

Pada proses *K-Means* yang akan dilakukan adalah menentukan nilai K yang ingin digunakan terlebih dahulu. Untuk mendapatkan K yang ideal dalam proses *K-Means* adalah dengan cara memasukan operator *Performance* dilihat dari nilai *Davies Bouldin Index* (DBI) yang paling kecil yaitu nilai yang mendekati 0. Proses menentukan nilai K yang ideal dapat dilihat pada Gambar 6.



Gambar 5. Proses menentukan nilai K ideal

Setelah itu jalankan proses operator performance, kemudian menghasilkan beberapa nilai DBI yang dapat dilihat pada Tabel 3. berikut.

Tabel 3. Hasil nilai K ideal

Nilai K	Max run	Davies Bouldin Index
2	10	0.131
3	10	0.132
4	10	0.116
5	10	0.163
6	10	0.101
7	10	0.087
8	10	0.097
9	10	0.110
10	10	0.109

Setelah dilakukan pengujian K= 2-10, nilai yang terkecil dibandingkan dengan nilai K yang lain diperoleh oleh K=7 dengan nilai DBI 0.087. Kemudian, nilai K 7 lah yang diambil untuk proses K-Means.

5. KESIMPULAN DAN SARAN

Berdasarkan pengolahan data yang telah dilakukan dapat diambil kesimpulan bahwa: Pengelompokan data bencana alam menggunakan algoritma K-Means clustering dengan bantuan Rapidminer, menghasilkan 7 cluster setelah dilakukan proses olah data. Cluster tersebut diantaranya yaitu, Cluster 6 dengan frekuensi sangat tinggi untuk kejadian bencana alam, yang terbagi menjadi 34 desa/kelurahan. Cluster 5 adalah cluster dengan frekuensi tinggi 1 untuk kejadian bencana alam nya, ada 2 desa/kelurahan pada cluster tersebut. Cluster 1, adalah cluster dengan kejadian bencana alam frekuensi sedang 1, memiliki 3 anggota desa/kelurahan. Cluster 3, adalah cluster kejadian bencana alam dengan frekuensi sedang 2, yang hanya terdapat 1 anggota desa/kelurahan. Cluster 0, adalah cluster yang memiliki frekuensi kejadian bencana alam sedang 3, yang memiliki 59 anggota desa/kelurahan. Cluster 4, adalah cluster frekuensi rendah 2 kejadian bencana alam nya, yaitu memiliki 1 anggota desa/kelurahan. Cluster 2, yaitu cluster dengan frekuensi kejadian bencana alam rendah 3 yang terdiri dari 10 anggota desa/kelurahan. Hasil

Pengujian K yang paling ideal pada proses Algoritma K-Means adalah K=7, dengan nilai Davies Bouldin Index (DBI) terkecil yaitu 0.087.

DAFTAR PUSTAKA

[1] Y. Pratama, H. Hendrawan, E. Rasywir, B. T. Carenina, and D. R. Anggraini, "Penerapan Algoritma K-Means clustering Untuk Mengelompokkan Provinsi Berdasarkan Banyaknya Desa/Kelurahan Dengan Upaya Antisipasi/Mitigasi Bencana Alam," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2549.

[2] W. Isworo Nugroho, "3 rd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) 30 Agustus 2023-Jakarta," 2023.

[3] W. Sri Utami, N. Pratiwi, and M. Faisal, "Bulletin of Information Technology (BIT) Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Clustering Perokok Usia Lebih dari 15 Tahun," vol. 4, no. 4, pp. 501–507, 2023, doi: 10.47065/bit.v3i1.

[4] A. Baldah, V. Duarisah, and R. A. Maulana, "Clustering Daerah Rawan Bencana Alam Di Indonesia Berdasarkan Provinsi Dengan Metode K-Means", [Online]. Available: <https://www.bps.go.id>

[5] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 2, pp. 156–165, Aug. 2022, doi: 10.23887/janapati.v11i2.45376.

[6] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 17–24, Apr. 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.

[7] L. Perdana Naibaho, A. Perdana Windarto, and K. Kunci, "Jurnal JPILKOM (Jurnal Penelitian Ilmu Komputer) PENERAPAN DATA MINING PADA TINGKAT PENGHUNIAN KAMAR PADA HOTEL BERBINTANG BERDASARKAN PROVINSI MENGGUNAKAN K-MEANS CLUSTERING Article Info Clustering Data Mining Hotel Kamar K-Means Keywords: Clustering Data Mining Hotels Rooms K-Means," Online, 2023.

[8] R. Ayuna Sari, H. Bayu Seta, and I. Gede Susrama Mas Diyasa, "Implementasi K-Means Clustering dengan Menggunakan Data Transaksi Penjualan untuk Penentuan Reward pada Agen Aqua dan Gas LPG FF Tirta," *JURNAL INFORMATIK Edisi ke*, vol. 18, p. 2022.

[9] M. R. Nahjan, N. Heryana, and A. Voutama, "IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK

ANALISA PENJUALAN PADA TOKO OJ CELL,” 2023.

- [10] Y. Yusfrizal, A. Meizar, N. Nurhayati, and H. Kurniawan, “Pengolahan Data Nilai Siswa Menggunakan Microsoft Excel,” *Zadama: Jurnal Pengabdian Masyarakat*, vol. 1, no. 1, pp. 85–91, Aug. 2022, doi: 10.56248/zadama.v1i1.25.