

IMPLEMENTASI ALGORITMA *K-MEANS* PADA PEMETAAN DAERAH TERDAMPAK TANAH LONGSOR DI JAWA TENGAH

Erika Rahmawati, Betha Nurina Sari, Mohamad Jajuli

Informatika, Universitas Singaperbangsa Karawang

Jl. HS. Ronggowaluyo, Telukjambe Timur, Karawang - 41363

erika.rahmawati18153@student.unsika.ac.id

ABSTRAK

Karakteristik wilayah dan letak geografis Indonesia menjadikan wilayah Indonesia menjadi daerah rawan bencana tanah longsor. Salah satu wilayah yang berpotensi tinggi mengalami bencana tanah longsor adalah Jawa Tengah. Dampak dari bencana tanah longsor di Jawa Tengah mengancam kehidupan masyarakat karena mengakibatkan rusaknya infrastruktur, merugikan harta benda dan menghilangkan nyawa manusia. Untuk meminimalisir dampak yang terjadi, diperlukan mitigasi bencana tanah longsor. Salah satu langkah awal dalam mitigasi bencana tanah longsor adalah pemetaan daerah. Pada penelitian ini dilakukan pengklasteran daerah berdasarkan dampak bencana tanah longsor di Jawa Tengah menggunakan algoritma *K-Means* dan pemetaan menggunakan QGIS. Dari metode *elbow* diperoleh nilai *k* yang paling optimal adalah $K=2$. Dari 2 *Cluster* yang terbentuk, sebanyak 24 daerah dengan kategori rendah dan sebanyak 7 daerah dengan kategori tinggi. Hasil pemodelan tersebut dievaluasi menggunakan *Silhouette Index* (SI) untuk mengetahui kualitas dari *cluster*. Dari hasil pengujian didapatkan nilai SI 0,48 dengan struktur lemah. Penelitian ini diharapkan dapat memberikan penyajian data pemetaan daerah terdampak tanah longsor di Jawa Tengah agar dapat digunakan sebagai informasi untuk meningkatkan kewaspadaan dan kesiapsiagaan akan terjadinya bencana tanah longsor.

Kata kunci : Algoritma *K-Means*, *Clustering*, Mitigasi Bencana, Tanah Longsor.

1. PENDAHULUAN

Wilayah Indonesia sangat rawan terhadap bencana tanah longsor antara lain karena karakteristik wilayah Indonesia yang terdiri dari dataran tinggi dan rendah dan memiliki curah hujan yang tinggi. Salah satu wilayah di Indonesia yang sering berpotensi mengalami bencana tanah longsor adalah provinsi Jawa Tengah. Menurut data BNPB, selama pertengahan awal tahun 2023, bencana tanah longsor paling banyak terjadi di Jawa Tengah. Dalam wawancara Kepala Bidang Penanganan Darurat BPBD Jawa Tengah, mengatakan bahwa kejadian bencana alam yang paling mendominasi di Jawa Tengah selama tahun 2022 adalah bencana tanah longsor[1]. Berdasarkan analisis dari penelitian sebelumnya, faktor pemicu terjadinya bencana tanah longsor di Jawa Tengah antara lain karena kondisi curah hujan dengan intensitas tinggi, topografi, dan penggunaan lahan[2].

Dampak yang ditimbulkan dari bencana tanah longsor di Jawa Tengah selama tahun 2021 hingga 2022 mengakibatkan rusaknya infrastruktur, merugikan harta benda, dan menghilangkan nyawa manusia. Hal ini membuktikan adanya bahaya dan mengancam kehidupan masyarakat.

Mitigasi bencana merupakan serangkaian upaya yang dilakukan untuk mengurangi resiko dari terjadinya suatu bencana dengan cara pembangunan fisik maupun penyadaran atau peningkatan dalam menghadapi ancaman bencana. Salah satu langkah untuk mengurangi akibat bencana tanah longsor yaitu penyusunan peta daerah rawan tanah longsor[3]. Selain pemetaan daerah rawan, beberapa langkah awal yang dapat dilakukan dalam mitigasi bencana tanah

longsor antara lain pembuatan prediksi, Pendidikan dan Latihan, relokasi pemukiman penduduk, pembuatan aturan konstruksi, pembuatan jalur dan rambu evakuasi, pembentukan satuan tugas, dan persiapan peralatan[4]. Sebagai informasi awal untuk memberikan kesadaran dan kesiapsiagaan dini bagi masyarakat diperlukan pemetaan daerah terdampak bencana tanah longsor. Pembuatan prediksi dapat dilakukan karena kejadian dan pola bencana tanah longsor memiliki pola kecenderungan berulang. Penelitian ini diharapkan dapat menjadi informasi sebagai langkah mitigasi awal.

2. TINJAUAN PUSTAKA

2.1. Algoritma *K-Means*

Algoritma *K-Means* merupakan salah satu algoritma dalam teknik *clustering* non hierarki yang membagi data ke dalam bentuk satu atau lebih *cluster*[5]. Algoritma *K-Means* mengelompokkan data dengan karakteristik yang sama ke dalam satu *cluster*. Langkah pertama algoritma *K-Means* menentukan jumlah *cluster* dan menginisiasi titik pusat awal. Langkah kedua yaitu masukkan semua data ke *cluster* terdekat. Kedekatan objek ditentukan oleh jarak menggunakan teori jarak *Euclidean* dengan rumus sebagai berikut[6]:

$$d(i,j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

Keterangan:

$d(i,j)$: jarak objek ke-*i* ke pusat *cluster* *j*

x_{ki} : data ke-*i* pada atribut data ke-*k*

x_{kj} : titik pusat ke-*j* pada atribut ke-*k*

Langkah ketiga yaitu hitung ulang titik pusat menggunakan rata-rata data dalam satu *cluster*. Dan langkah terakhir yaitu mengulangi langkah kedua dan ketiga sampai konvergen, yaitu sampai tidak ada lagi perpindahan objek dari satu *cluster* ke *cluster* yang lain.

2.2. Metode Elbow

Metode *elbow* merupakan metode yang berguna untuk menentukan hasil *cluster* terbaik. Jika terdapat nilai yang membentuk siku dalam grafik atau memiliki selisih nilai paling besar, maka jumlah *cluster* tersebut merupakan *cluster* yang paling optimal[7].

2.3. Silhouette Index (SI)

Silhouette Index (SI) merupakan salah satu teknik validasi atau penguji kualitas *cluster*. *Silhouette Index* merupakan integrasi dari metode validasi *clustering* yaitu kohesi yang berfungsi untuk mengetahui seberapa dekat jarak dalam *cluster*, dan separasi. yang berfungsi untuk mengetahui seberapa jauh jarak antar *cluster*. Hasil perhitungan nilai SI menunjukkan ketepatan dalam pengelompokkan data. Hasil nilai SI bervariasi dari rentang -1 sampai dengan 1.

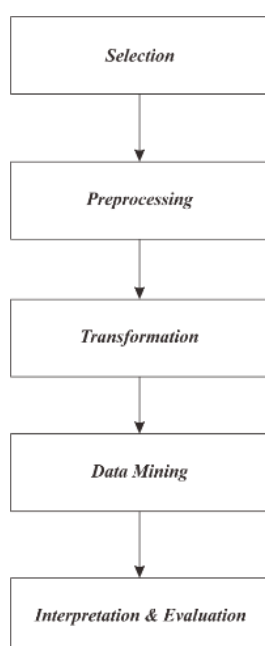
Tabel 1. Kategori nilai *Silhouette Index*

No	Rentang Nilai SI	Kategori
1	$0,7 < SI \leq 1$	<i>Strong Structure</i>
2	$0,5 < SI \leq 0,7$	<i>Medium Structure</i>
3	$0,25 < SI \leq 0,5$	<i>Weak Structure</i>
4	$SI \leq 0,25$	<i>No Structure</i>

Tabel 1 merupakan rincian predikat kategori nilai *Silhouette Index* menurut Kaufman dan Rousseeuw[8].

3. METODE PENELITIAN

3.1. Alur Penelitian



Gambar 1. Alur Penelitian

Diagram alir pada Gambar 1 merupakan metodologi yang digunakan dalam penelitian ini yaitu *Knowledge Discovery in Database* (KDD). KDD adalah salah satu metodologi penelitian yang berguna dalam proses *data mining*. KDD bertujuan untuk menemukan dan mengidentifikasi suatu pola dalam *database* untuk mendapatkan informasi baru. Tahapan KDD terdiri dari *selection*, *preprocessing*, *transformation*, *data mining*, *interpretation & evaluation*[5].

3.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder. Data yang digunakan merupakan data tanah longsor di Provinsi Jawa Tengah. Dataset diambil dari situs cewadis.bpbdd.jatengprov.go.id milik Badan Penanggulangan Bencana Daerah Jawa Tengah. Data diunduh pada tanggal 7 Desember 2023 pukul 08.52.

3.3. Selection

Pada tahap ini dilakukan seleksi pada dataset. Provinsi Jawa Tengah memiliki 35 Kabupaten/Kota, diseleksi menjadi 31 record dan 31 atribut.

3.4. Preprocessing

Pada tahap ini dilakukan *data cleaning*, yaitu pengecekan *missing value*, pengecekan duplikasi data, dan penghapusan atribut yang tidak sesuai dengan tujuan *data mining*.

3.5. Transformation

Dalam tahap ini dilakukan normalisasi data dengan teknik standarisasi dan reduksi data dengan mereduksi dimensi menggunakan teknik *Principal Component Analysis* (PCA). Data harus distandarisasi sebelum penerapan PCA agar interval data menjadi lebih proporsional [9]. PCA dilakukan dengan tujuan untuk meningkatkan kualitas hasil *clustering* [10].

3.6. Data Mining

Langkah berikutnya yaitu *clustering* dilakukan dengan menggunakan algoritma *K-Means* untuk mengelompokkan [11]. Untuk mendapatkan nilai k yang optimal didapatkan dari metode *elbow* [12].

3.7. Interpretation & Evaluation

Pada tahap terakhir hasil *cluster* yang paling optimal akan diinterpretasikan dan divisualisasikan dalam bentuk peta tematik, kemudian dievaluasi menggunakan *Silhouette Index*.

4. HASIL DAN PEMBAHASAN

4.1. Selection

Pada tahap pertama penelitian ini dilakukan pemilihan data agar relevan dengan tujuan *data mining*.

Tabel 2. Dataset Hasil Seleksi

Kode Wilayah	Nama Wilayah	Jumlah Kejadian	...	Taksiran Kerugian
3301	Cilacap	128	...	1.511.833.500
3302	Banyumas	142	...	141.500.000
...
...
3376	Kota Tegal	1	...	2.000.000

Tabel 2 merupakan beberapa contoh data hasil seleksi dari situs cevadis.bppd.jatengprov.go.id. Adapun atribut-atribut yang digunakan antara lain.

Tabel 3. Atribut Hasil Seleksi

No	Atribut	Tipe Data
1	Kode Wilayah	Kategorik
2	Nama Wilayah	Kategorik
3	Jumlah Kejadian	Numerik
4	Tanah Longsor	Numerik
5	Meninggal	Numerik
6	Hilang	Numerik
7	Luka Ringan	Numerik
8	Luka Berat	Numerik
9	Pengungsi-KK	Numerik
10	Pengungsi-Jiwa	Numerik
11	Ancaman-KK	Numerik
12	Ancaman-Jiwa	Numerik
13	RMH-RB	Numerik
14	RMH-RS	Numerik
15	RMH-RR	Numerik
16	RMH Terancam	Numerik
17	RMH Terendam	Numerik
18	Sekolah	Numerik
19	Tempat Ibadah	Numerik
20	Kios	Numerik
21	Bangunan Lain	Numerik
22	Jembatan	Numerik
23	Jalan	Numerik
24	Sawah	Numerik
25	Kebun-Hutan	Numerik
26	Kolam	Numerik
27	Hewan Ternak	Numerik
28	Talud	Numerik
29	Tanggul	Numerik
30	Irigasi	Numerik
31	Taksiran Kerugian	Numerik

Dari Tabel 3 dapat dilihat dengan rinci 31 atribut yang akan digunakan pada penelitian.

4.2. Preprocessing

Tahap selanjutnya merupakan tahap *preprocessing*. Pada tahap ini akan dilakukan pembersihan data. Dari 31 atribut sebelumnya akan diseleksi apakah terdapat *missing value* atau tidak. Data dikatakan *missing value* apabila apabila atribut-atribut dalam dataset tidak berisi nilai atau kosong.



Gambar 2. Pengecekan *missing value* pada tools RapidMiner

Dapat diketahui dari Gambar 2 bahwa tidak terdapat *missing value* pada dataset. Karena tidak terdapat *missing value* maka dapat dilanjutkan ke proses berikutnya, yaitu menghapus duplikasi data.

Jumlah Kejadian	Tanah Longsor
128	128
142	142
32	32
42	42
27	27

Gambar 3. Contoh Duplikasi Atribut

Pada Gambar 3 menunjukkan bahwa ada duplikasi atribut dalam dataset. Pada atribut Jumlah Kejadian dan atribut Tanah Longsor memiliki *value* yang sama. Agar tidak terjadi duplikasi data, atribut Tanah Longsor akan dihapus. Selain duplikasi data, atribut yang tidak dibutuhkan seperti atribut Kode Wilayah juga akan dihapus.

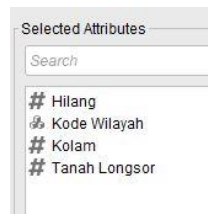
name	Type	Missing	Statistics	Filter (1/1 attribute)	Search for attributes
Hilang	Integer	0	Min: 0	Max: 0	Average: 0

Gambar 4. Statistik pada Atribut Hilang

name	Type	Missing	Statistics	Filter (1/1 attribute)	Search for attributes
Kolam	Integer	0	Min: 0	Max: 0	Average: 0

Gambar 5. Statistik pada Atribut Kolam

Gambar 4 dan Gambar 5 menunjukkan bahwa seluruh *record* pada atribut Hilang dan atribut Kolam bernilai 0 atau berarti tidak terdampak. Hal ini berarti tidak ada kasus jumlah orang hilang dan kolam terdampak pada kejadian tanah longsor. Maka atribut Hilang dan atribut Kolam akan dihapus karena tidak sesuai dengan tujuan *clustering*.



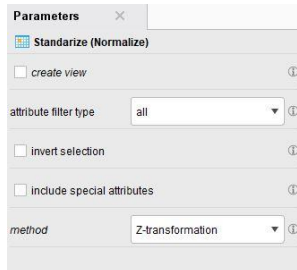
Gambar 6. Atribut yang akan dihapus

Dapat dilihat dari Gambar 6 merupakan proses penghapusan seluruh atribut yang tidak relevan dengan tujuan *clustering* yaitu atribut Tanah Longsor, atribut Kode Wilayah, atribut Hilang, dan atribut Kolam yang

dilakukan pada *tools RapidMiner* menggunakan fitur *Selected Attributes*.

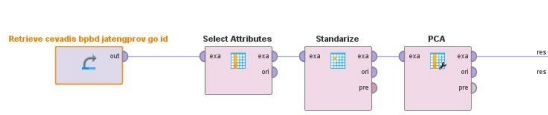
4.3. Transformation

Setelah data bersih, tahap berikutnya dilakukan transformasi data. Pada tahap ini dilakukan standarisasi dan reduksi data.



Gambar 7. Parameter Standarisasi Data pada *tools RapidMiner*

Gambar 7 adalah proses transformasi data menggunakan teknik standarisasi dengan *tools RapidMiner*. Setelah data proporsional, tahap selanjutnya dilakukan reduksi data. Reduksi data merupakan proses pengurangan dimensi data asli dan merepresentasikannya dalam dimensi yang lebih kecil tanpa kehilangan informasi yang ada. Salah satu teknik dalam mereduksi data adalah PCA. PCA dapat menghilangkan karakteristik yang tidak relevan dan juga dapat mereduksi dimensi data pada data besar tanpa menghilangkan seluruh informasi.



Gambar 8. Proses PCA

Gambar 8 merupakan proses penerapan PCA pada *RapidMiner*. PCA membentuk atribut baru bernama PC. Atribut PC ini akan menggantikan atribut awal pada analisis selanjutnya. Atribut baru PC merupakan kombinasi linier dari atribut awal.

Tabel 4. Nilai *Eigen* Hasil PCA

Komponen	Total	%Varian	%Kumulatif
PC 1	2,895	0,322	0,322
PC 2	1,943	0,145	0,468
PC 3	1,578	0,096	0,563
PC 4	1,441	0,080	0,643
PC 5	1,311	0,066	0,709
PC 6	1,278	0,063	0,772
PC 7	1,188	0,054	0,826
PC 8	1,101	0,047	0,873
PC 9	0,948	0,035	0,908
PC 10	0,892	0,031	0,938
PC 11	0,747	0,021	0,960
PC 12	0,616	0,015	0,974
PC 13	0,499	0,010	0,984
PC 14	0,386	0,006	0,990
PC 15	0,322	0,004	0,994
PC 16	0,286	0,003	0,997
PC 17	0,173	0,001	0,998
PC 18	0,168	0,001	0,999
PC 19	0,104	0,000	0,999
PC 20	0,087	0,000	1,000
PC 21	0,066	0,000	1,000
PC 22	0,039	0,000	1,000
PC 23	0,028	0,000	1,000
PC 24	0,021	0,000	1,000
PC 25	0,017	0,000	1,000
PC 26	0,012	0,000	1,000

Tabel 4 menyajikan dengan rinci nilai *eigen* yang menunjukkan jumlah varian yang mampu dijelaskan oleh suatu komponen utama. PC 1 adalah atribut yang menjelaskan keragaman paling besar diikuti PC 2, PC 3, dan seterusnya. Dapat dilihat dari Tabel 3 diperoleh 8 *principal component* yang dilihat berdasarkan nilai *eigen* > 1 yang layak dipilih [10]. Dari 8 *principal component* tersebut, sebesar 87,3% varians dari 26 atribut dapat dijelaskan oleh *principal component* 1 sampai 8.

Tabel 5. Hasil *Loding Value* PCA

Atribut	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8
Jumlah Kejadian	0,329	0,073	0,011	-0,090	0,037	0,020	0,109	-0,038
Meninggal	0,075	0,267	0,158	-0,036	-0,237	0,198	-0,400	0,135
Luka Ringan	0,086	0,202	0,224	-0,166	0,016	0,330	-0,479	-0,102
Luka Berat	0,021	-0,057	0,459	-0,092	0,293	0,050	0,129	0,325
Pengungsi-KK	0,287	0,197	-0,020	0,069	-0,202	-0,180	0,071	0,017
Pengungsi-Jiwa	0,307	0,138	0,041	-0,100	-0,012	-0,190	0,037	0,035
Ancaman-KK	0,175	0,020	-0,295	0,293	0,314	0,303	0,065	0,067
Ancaman-Jiwa	0,120	0,041	-0,302	0,331	0,318	0,298	0,113	0,066
RMH-RB	0,137	-0,279	0,314	0,235	0,069	-0,132	-0,143	-0,119
RMH-RS	0,089	-0,231	0,386	0,367	0,044	-0,051	-0,066	-0,083
RMH-RR	0,302	0,002	0,172	0,159	0,007	-0,125	0,029	0,022
RMH Terancam	0,312	0,035	-0,183	0,056	0,024	-0,153	-0,077	0,067
RMH Terendam	0,239	0,264	0,057	0,065	-0,168	-0,224	0,205	-0,044
Sekolah	0,147	-0,389	0,074	0,052	-0,135	0,080	-0,035	0,073
Tempat Ibadah	0,129	-0,288	0,136	0,233	-0,229	0,242	0,156	-0,016
Kios	0,213	-0,016	0,128	-0,372	0,310	0,044	0,149	-0,225

Atribut	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8
Bangunan Lain	0,103	-0,078	0,028	-0,185	-0,427	0,378	0,342	0,113
Jembatan	0,093	-0,306	-0,134	-0,036	-0,058	-0,041	-0,142	-0,107
Jalan	0,283	-0,099	-0,200	-0,038	0,150	0,016	-0,229	0,055
Sawah	-0,038	0,027	0,109	-0,092	0,165	-0,119	0,062	0,754
Kebun-Hutan	0,159	-0,230	-0,196	-0,123	0,041	-0,237	-0,339	0,230
Hewan Ternak	0,131	0,214	0,080	0,127	-0,018	0,323	-0,254	0,064
Talud	0,197	-0,035	0,132	-0,355	0,323	0,127	0,143	-0,259
Tanggul	0,176	-0,195	-0,175	-0,206	-0,154	-0,128	-0,131	-0,044
Irigasi	0,266	0,244	0,024	0,133	-0,148	-0,024	0,152	0,047
Taksiran Kerugian	0,329	0,073	0,011	-0,090	0,037	0,020	0,109	-0,038

Dari Tabel 5 dapat dilihat untuk menentukan atribut yang terbentuk ke dalam 8 *principal component*. Nilai *loading* dengan huruf tebal artinya terdapat korelasi antara atribut dengan *principal component* yang terbentuk.

Tabel 6. Hasil *Principal Component*

Atribut	<i>Principal Component (PC)</i>
Jumlah Kejadian, Pengungsi-KK, Pengungsi-Jiwa, RMH-RR, RMH Terancam, Jalan, Irigasi, Taksiran Kerugian	PC 1
RMH Terendam, Sekolah, Tempat Ibadah, Jembatan	PC 2
Luka Berat, RMH-RB, RMH-RS	PC 3
Ancaman-Jiwa, Kios, Talud, Tanggul	PC 4
Ancaman-KK, Bangunan Lain	PC 5
Hewan Ternak	PC 6
Meninggal, Luka Ringan, Kebun-Hutan	PC 7
Sawah	PC 8

Pada Tabel 6 didapatkan posisi masing-masing komponen membentuk 8 *principal component* dari nilai *loading* pada Tabel 4.

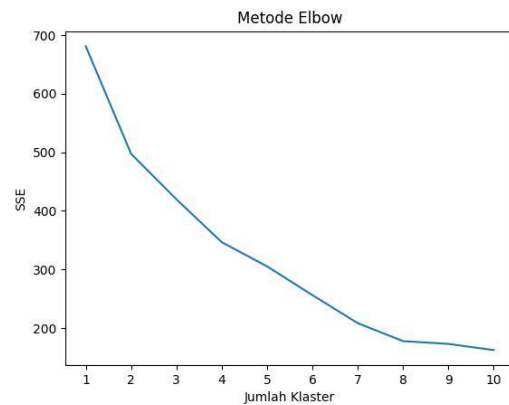
4.4. Data Mining

Data mining merupakan proses penerapan model untuk mencapai tujuan *clustering*. Pada tahap ini akan dilakukan *clustering* menggunakan algoritma *k-means* dengan rumus jarak *euclidean distance*. Jumlah *cluster* ditentukan dengan metode *elbow* menggunakan nilai *Sum of Square Error (SSE)*.

Tabel 7. Hasil Selisih SSE

Klaster	Nilai SSE	Selisih
K = 1	680,9936525806451	0
K = 2	497,39397778571424	183,5997
K = 3	419,85114883333324	77,54283
K = 4	346,5309551818182	73,32019
K = 5	305,32918110606056	41,20177
K = 6	256,2820961315789	49,04708
K = 7	208,41188438157894	47,87021
K = 8	177,734014	30,67787
K = 9	172,99773575	4,736278
K = 10	162,45563928787874	10,5421

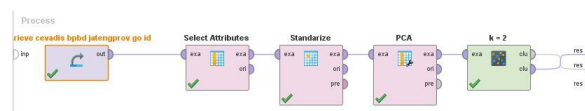
Tabel 7 merupakan hasil perbandingan nilai SSE yang telah dilakukan pada K=1 sampai K=10. Selisih paling besar terjadi pada saat K=2.



Gambar 9. Grafik Metode *Elbow*

Dapat dilihat dari Gambar 9 grafik membentuk siku pada K=2. Maka dapat disimpulkan bahwa nilai K yang paling optimal adalah pada saat nilai SSE memiliki selisih paling besar dan grafik menunjukkan penurunan paling signifikan, yaitu pada K=2.

Selanjutnya dilakukan pengujian kembali pada *RapidMiner* menggunakan *clustering* yang direkomendasikan, yaitu dengan parameter K=2.



Gambar 10. Proses *Clustering*

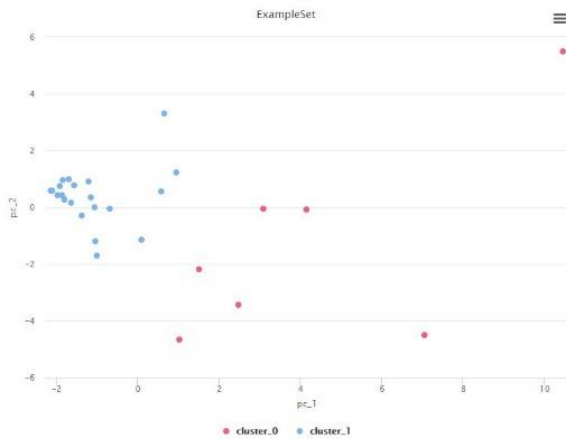
Gambar 10 merupakan proses *clustering* dengan algoritma *K-Means* menggunakan tools *RapidMiner*.

Cluster Model

```
Cluster 0: 7 items
Cluster 1: 24 items
Total number of items: 31
```

Gambar 11. *Cluster Model*

Dari hasil *clustering* didapatkan model seperti pada Gambar 11. Terdapat 7 anggota Cluster 0 dan 24 anggota Cluster 1.



Gambar 12. Visualisasi cluster pada tools RapidMiner

Pada tools RapidMiner Cluster 0 ditunjukkan dengan warna merah dan Cluster 1 ditunjukkan dengan warna biru seperti pada Gambar 12.

4.5. Interpretation & Evaluation

Setelah dilakukan analisis pada cluster dengan menggunakan algoritma k-means tahap selanjutnya dilakukan interpretasi. Interpretasi dari hasil clustering menunjukkan adanya perbedaan karakteristik dari Cluster 0 dan Cluster 1, yaitu dapat dilihat pada data rata-rata setiap atribut pada dataset dampak bencana tanah longsor. Cluster 0 menunjukkan rata-rata jumlah kejadian bencana longsor sebanyak 122 kali dalam 3 tahun, sedangkan Cluster 1 hanya 17 kali dalam 3 tahun. Adapun rincian karakteristik dari 2 cluster dilihat dari rata-rata 31 atribut dapat dilihat pada Tabel 8.

Tabel 8. Karakteristik Cluster

Atribut	Cluster 0	Cluster 1
Jumlah Kejadian	122,57	17,25
Meninggal	0,57	0,58
Luka Ringan	1,43	0,75
Luka Berat	0,29	0,13
Pengungsi-KK	17,57	1,13
Pengungsi-Jiwa	82,14	5,92
Ancaman-KK	7,57	1,83
Ancaman-Jiwa	21,71	9
RMH-RB	11,29	1,83
RMH-RS	22,29	2,21
RMH-RR	60,57	9,67
RMH Terancam	65,57	6,71
RMH Terendam	17,14	0
Sekolah	2,43	0,33
Tempat Ibadah	2	0,21
Kios	4,14	0,25
Bangunan Lain	2,43	0,58
Jembatan	3	0,83
Jalan	8,29	1,42
Sawah	0,43	2,5
Kebun-Hutan	3,14	0
Hewan Ternak	0,43	0,21
Talud	29,14	3,54
Tanggul	2,57	0,25
Irigasi	2,57	0,25

Atribut	Cluster 0	Cluster 1
Taksiran Kerugian	636586785,7	176811895,8
Jumlah Kejadian	11,29	17,25
Meninggal	22,29	0,58
Luka Ringan	60,57	0,75
Luka Berat	65,57	0,13
Pengungsi-KK	17,14	1,13

Berdasarkan Tabel 8 dapat diketahui rata-rata nilai dari anggota Cluster 0 lebih besar dari Cluster 1. Maka dapat disimpulkan bahwa Cluster 0 merupakan kategori tinggi dan Cluster 1 merupakan kategori rendah. Dengan rincian anggota sebagai berikut.

Tabel 9. Anggota Cluster

Cluster	Anggota	Jumlah	Kriteria
Cluster 0	Cilacap, Banyumas, Magelang, Temanggung, Tegal, Brebes, Kota Semarang.	7	Tinggi
Cluster 1	Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Boyolali, Klaten, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembang, Pati, Kudus, Jepara, Semarang, Kendal, Batang, Pekalongan, Pemalang, Kota Magelang, Surakarta, Kota Tegal	24	Rendah

Tabel 9 menunjukkan rincian bahwa diperoleh 2 cluster daerah terdampak tanah longsor di Jawa Tengah, yaitu Cluster 0 yang berjumlah 7 Kabupaten/Kota dengan tingkat tinggi dan Cluster 1 yang berjumlah 24 Kabupaten/Kota dengan tingkat rendah. Adapun hasil clustering ini divisualisasikan dalam bentuk peta menggunakan QGIS, yang dapat dilihat pada Gambar 13.



Gambar 13. Peta Cluster Dampak Tanah Longsor di Jawa Tengah 2020-2022

Gambar 13 merupakan visualisasi daerah terdampak tanah longsor di Jawa Tengah tahun 2020 – 2022 dalam bentuk pemetaan. Daerah terdampak tanah longsor di Jawa Tengah tahun 2020-2022 terbagi menjadi 2 *cluster*. Daerah dengan warna hijau merupakan Cluster 1 yang berarti daerah dengan dampak tanah longsor tingkat rendah, sedangkan daerah dengan warna merah merupakan Cluster 0 yang berarti daerah dengan dampak tinggi. Daerah dengan warna putih tidak termasuk ke dalam *cluster* manapun karena tidak terdampak tanah longsor.

Pada tahap terakhir hasil *cluster* yang paling optimal akan dievaluasi menggunakan *Silhouette Index*. Hasil pengujian menggunakan *silhouette index* atau *silhouette score* menghasilkan nilai indeks sebesar 0,48 pada hasil model *clustering* dengan $K=2$. Kualitas *cluster* yang dihasilkan dari model *clustering* ini termasuk dalam kategori *weak structure* [13].

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang sudah dilakukan, maka kesimpulan yang didapatkan adalah Algoritma K-means mampu mengetahui pengelompokan daerah tanah longsor di Jawa Tengah berdasarkan dampak-dampak yang diakibatkan. Pada pengelompokan daerah terdampak tanah longsor di Jawa Tengah pada tahun 2020-2022 didapatkan sebanyak 24 wilayah masuk ke dalam *cluster* tingkat rendah dan 7 wilayah masuk ke dalam *cluster* tingkat tinggi. Hasil visualisasi daerah terdampak tanah longsor di Jawa Tengah tahun 2020-2022 menggunakan QGIS didapatkan area dengan warna merah merupakan Cluster 0 yang berarti daerah dengan dampak tinggi, sedangkan area dengan warna hijau merupakan Cluster 1 yang berarti daerah dengan dampak rendah. Area dengan warna putih yang terdiri dari Kabupaten Demak, Kabupaten Sukoharjo, Kota Pekalongan, dan Kota Salatiga tidak termasuk ke dalam *clustering* karena tidak terdampak dalam bencana tanah longsor.

Penelitian ini diharapkan dapat memberikan penyajian data pemetaan daerah terdampak tanah longsor di Jawa Tengah agar dapat digunakan sebagai informasi untuk meningkatkan kewaspadaan dan kesiapsiagaan akan terjadinya bencana tanah longsor. Dari hasil penelitian yang telah dilakukan, saran yang diberikan untuk menjadi referensi dalam penelitian selanjutnya diharapkan mengintegrasikan hasil pemetaan pada sebuah *website*.

DAFTAR PUSTAKA

- [1] P. Utomo, "Bencana Alam Menimpa Jateng Sepanjang 2022, Longsor Mendominasi," *RRI*, Semarang, 2023. [Online]. Available: <https://rri.co.id/daerah/145085/bencana-alam-menimpa-jateng-sepanjang-2022-longsor-mendominasi>
- [2] R. Isnaini, "Analisis Bencana Tanah Longsor di Wilayah Provinsi Jawa Tengah," *Islam. Manag. Empower. J.*, vol. 1, no. 2, pp. 143–160, 2019, doi: 10.18326/imej.v1i2.143-160.
- [3] S. Mintarjo, *Waspada Tanah Longsor*. Bandung, 2018.
- [4] P. Supriyono, *Seri Pendidikan Pengurangan Risiko Bencana Tanah Longsor*. Yogyakarta: ANDI, 2014.
- [5] D. Jollyta, W. Ramdhan, and M. Zarlis, *Konsep Data Mining Dan Penerapan*. Sleman: Deepublish, 2020.
- [6] D. Jollyta, M. Siddik, H. Mawengkang, and S. Efendi, *Teknik Evaluasi Cluster Solusi Menggunakan Python dan RapidMiner*. Sleman: Deepublish, 2021.
- [7] R. Y. Sari *et al.*, "Algoritma K-Means dengan Metode Elbow untuk Mengelompokkan Kabupaten/Kota di Jawa Tengah Berdasarkan Komponen Pembentuk Indeks Pembangunan Manusia," vol. 3, no. 2, pp. 104–108, 2022.
- [8] M. Rahmah, A. Candra, and R. W. Sembiring, "Identifikasi Predikat Hasil Pengelompokan Data Kualitas Udara dengan Menggunakan Affinity Propagation dan Silhouette Coefficient," *InfoTekJar J. Nas. Inform. dan Teknol. Jar.*, vol. 6, no. 2, pp. 177–180, 2022, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/4670>
- [9] G. Enzellina and D. Suhaedi, "Penggunaan Metode Principal Component Analysis dalam Menentukan Faktor Dominan," *J. Ris. Mat.*, pp. 101–110, 2022, doi: 10.29313/jrm.v2i2.1192.
- [10] T. R. Yulianti, K. N. Siregar, A. Prabawa, and N. Fadhilah, "Identifikasi Atribut dengan Principal Component Analysis dan K-Means Clustering Sebagai Dasar Penyusunan Strategi Promosi KB Pria di Indonesia," *J. Biostat. Kependudukan, dan Inform. Kesehat.*, vol. 2, no. 2, p. 79, 2022, doi: 10.51181/bikfokes.v2i2.5868.
- [11] M. Herviany, S. P. Delima, T. Nurhidayah, and Kasini, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat," *Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021, [Online]. Available: <https://journal.irpi.or.id/index.php/malcom/article/view/60>
- [12] M. Orisa, "Optimasi Cluster pada Algoritma K-Means," *Pros. SENIATI*, vol. 6, no. 2, pp. 430–437, 2022, doi: 10.36040/seniati.v6i2.5034.
- [13] R. Hidayati, A. Zubair, A. H. Pratama, and L. Indana, "Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering," *Techno.Com*, vol. 20, no. 2, pp. 186–197, 2021, doi: 10.33633/tc.v20i2.4556.