



Analisis Metode *Clustering K-Means* pada Zonasi Daerah Terdampak Banjir di Kota Medan dengan Evaluasi *Silhouette Coefficient*

Febby Arisca Zurfani^{1*}, Sawaluddin², Mardiningsih³, Muhammad Romi Syahputra⁴

^{1,2,3,4}Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sumatera Utara, Indonesia

*Korespondensi penulis: febbyariscazurfani@gmail.com

Abstract. *Clustering is one of the fields of study that discusses data. Clustering is used to find and group data based on its traits or characteristics. Clustering can also be used for class-identified data. However, the clustering method automatically clusters the data before the class identifier is known. Based on the data obtained, the city of Medan, which has a population of approximately 2,460,858 people and an area of 26,510 hectares or 3.6% of the total area of North Sumatra Province, is classified as Flood-prone (BPS). Floods that occur almost 10 to 12 times a year in Medan City are influenced by the condition of the downstream Deli and Belawan rivers. Based on the results of the k-means clustering that has been carried out, the areas that are safe from flooding are the districts of Medan Amplas, Medan Denai, Medan Area, Medan Kota, Medan Petisah, Medan Perjuangan, Medan Tembung, Medan Deli, and Medan Labuhan. Areas prone to flooding are Medan Tuntungan, Medan Sunggal, Medan Helvetia, West Medan, and Medan Marelan. Meanwhile, the areas most prone to flooding are Medan Johor, Medan Maimun, Medan Polonia, Medan Baru, Medan Selayang, Medan Timur, and Medan Belawan based on the evaluation of the accuracy of the silhouette method of 0.9 and can be declared significant.*

Keywords: *Clustering, K-Means, Silhouette Coefficient.*

Abstrak. *Clustering adalah salah satu bidang studi yang membahas tentang data. Clustering digunakan untuk menemukan dan mengelompokkan data berdasarkan ciri-ciri atau karakteristiknya. Clustering juga dapat digunakan untuk data yang diberi pengenalan kelas. Namun, metode clustering secara otomatis mengelompokkan data sebelum pengidentifikasi kelas diketahui. Berdasarkan data yang diperoleh, Kota Medan yang berpenduduk kurang lebih 2.460.858 jiwa dan luas wilayah 26.510 hektar atau 3,6% dari total wilayah Provinsi Sumatera Utara tergolong Rawan Banjir (BPS). Banjir yang terjadi hampir 10 sampai 12 kali dalam setahun di Kota Medan dipengaruhi oleh kondisi hilir sungai Deli dan Belawan. Berdasarkan hasil clustering k-means yang telah dilakukan, adapun daerah yang aman dari bencana banjir yaitu kecamatan Medan Amplas, Medan Denai, Medan Area, Medan Kota, Medan Petisah, Medan Perjuangan, Medan Tembung, Medan Deli, dan Medan Labuhan. Daerah yang rawan banjir adalah Medan Tuntungan, Medan Sunggal, Medan Helvetia, Medan Barat, dan Medan Marelan. Sedangkan daerah yang paling rawan banjir adalah Medan Johor, Medan Maimun, Medan Polonia, Medan Baru, Medan Selayang, Medan Timur, dan Medan Belawan berdasarkan evaluasi keakuratan metode silhouette 0,9 dan dapat dinyatakan signifikan.*

Kata Kunci: *Clustering, K-means, Silhouette Coefficient.*

1. LATAR BELAKANG

Indonesia adalah satu dari negara-negara yang terletak di sepanjang garis khatulistiwa, yang memberikannya iklim dan musim yang berbeda. Indonesia punya dua periode cuaca, yakni musim hujan dan musim kering. Pada saat musim hujan tiba, daerah-daerah tertentu di Indonesia sering kali mengalami banjir yang menimbulkan dampak yang serius. Salah satu daerah yang tergolong dalam catatan langganan banjir adalah Kota Medan. Terdapat beberapa

catatan yang membuktikan bahwa tahun ke tahun Kota Medan ialah kota yang memiliki beberapa kecamatan yang patut diperhatikan Pemerintah.

Menurut Badan Nasional Penanggulangan Bencana (2022) banjir adalah suatu bencana alam yang menyebabkan suatu daerah terendam akibat volume air yang meningkat. Menurut informasi yang terdokumentasikan dalam Rencana Aksi Nasional Pengurangan Risiko Bencana (RAN-PRB), jumlah bencana yang terjadi di Indonesia setiap tahun terus bertambah. Akibatnya, masyarakat Indonesia mengalami kerugian yang cukup besar, baik kerugian ekonomi maupun fasilitas (Yana et al. 2018).

Salah satu bencana alam yang menjadi sorotan Indonesia adalah bencana banjir. Bencana banjir merupakan bencana yang mengakibatkan daerah terdampak banjir akan terendam yang diakibatkan oleh debit air yang meningkat secara signifikan sehingga menimbulkan berbagai macam kerugian, baik itu kerugian fisik maupun kerugian non fisik yang dapat menimbulkan korban seperti fasilitas masyarakat dan Pemerintah (bangunan sekolah, rumah sakit, rumah Peribadatan dan lain sebagainya) (Ravena, 2019).

Salah satu upaya yang dapat dilakukan sebagai bentuk tindakan preventif adalah mitigasi bencana. Mitigasi bencana ialah suatu tindakan yang dilakukan untuk meminimalisir dampak yang terjadi akibat dari bencana atau persiapan yang dilakukan untuk mengatasi resiko jangka panjang (Sari, 2018).

Daerah rawan banjir adalah wilayah yang rentan mengalami banjir, baik yang disebabkan oleh meluapnya sungai akibat curah hujan tinggi maupun oleh banjir bandang, yang merupakan banjir besar dan mendadak akibat turunnya hujan deras secara tiba-tiba dalam jangka waktu yang sangat singkat dan dapat menyebabkan longsor di hilir sungai dan bendungan alam (Novi, 2024).

Berdasarkan data yang diperoleh, Kota Medan yang berpenduduk kurang lebih 2.460.858 jiwa dan luas wilayah 26.510 hektar atau 3,6% dari total wilayah Provinsi Sumut merupakan daerah rawan banjir (BPS Utara). Sumatera (2022) Kota Medan mengalami banjir hampir 10-12 kali dalam setahun, hal ini dipengaruhi oleh keadaan sungai Deli dan Belawan di hulunya (Tampubolon, 2018).

Persoalan banjir yang terjadi di kota Medan sudah tergolong kategori penyakit kronis bahkan menjadi tradisi tahunan. Telah banyak upaya dilakukan untuk menanggulangnya bahkan tidak terhitung berapa besar dana yang dikeluarkan di berbagai proyek. Nyatanya dana tersebut bagaikan air di padang pasir, seluruh dana yang telah dicurahkan rakyat habis, sedangkan banjir selalu menjadi kegelisahan bagi masyarakat kota Medan. Hingga sekarang banjir terus meresahkan 2,1 juta masyarakat kota Medan. Hal tersebut dikarenakan banjir tidak

bergantung pada hujan yang turun di hulu sungai Deli, hujan di kota Medan dapat mengakibatkan orang Medan kewalahan karena genangan air yang ada di mana-mana. Sama halnya dengan kawasan permukiman yang padat penduduknya akan jadi langganan banjir, apalagi hujan deras turun di bagian hulu sungai yang melintasi kota Medan (Bhakti Alamsyah, 2018).

Zonasi merupakan istilah yang berasal dari zona yang menggambarkan wilayah atau area dengan ciri khas lingkungan yang spesifik dan memiliki fungsi tertentu. Zoning in English is known as Zonasi. Di beberapa negara, peraturan penggunaan lahan dikenal dengan nama yang berbeda, seperti rencana penggunaan lahan, peraturan penggunaan lahan, keputusan penggunaan lahan, undang-undang penggunaan lahan, undang-undang perencanaan kota, undang-undang zonasi, dan lain-lain (Ade Heryana, 2020). Adapun beberapa wilayah kota Medan berdasarkan catatan pemko Medan nomor 13 tahun 2013.

Clustering merupakan salah satu ilmu dalam bidang analisis data. Clustering merupakan teknik yang dapat digunakan dalam mengidentifikasi dan juga mengelompokkan data berdasarkan ciri cirinya dan mengidentifikasi pola yang ada didalamnya. Berbeda dari klasifikasi, teknik clustering mengelompokkan data secara otomatis tanpa perlu label kelas terlebih dahulu. Namun, data yang sudah dikelompokkan berdasarkan kelas yang sudah diketahui juga bisa diatur dalam kelompok (Suyanto, 2017).

Dalam aplikasi pengelompokan data, clustering memiliki peran krusial, seperti dalam eksplorasi data, analisis web, analisis teks, dan aplikasi basis data yang berfokus pada lokasi. Penelitian ini menggunakan teknik clustering untuk mengidentifikasi wilayah yang rentan terhadap bencana.

Menurut Santoso (2002), karakteristik dari sebuah cluster adalah sebagai berikut:

- 1) Memiliki tingkat homogenitas yang tinggi di antara anggotanya (within cluster), dan
- 2) Memiliki perbedaan yang signifikan antara cluster-cluster (between cluster).

Salah satu metode yang umum digunakan untuk clustering adalah metode K-Means. Metode K-Means bertujuan untuk membagi data ke dalam beberapa cluster sedemikian rupa sehingga varians internal (error) di dalam setiap cluster minimal. Algoritma K-Means dimulai dengan menentukan titik-titik pusat awal dari setiap cluster. Proses utama K-Means adalah menghitung jarak Euclidean kuadrat antara setiap data masukan dengan pusat cluster untuk menentukan keanggotaan data tersebut ke salah satu cluster. Setelah itu, pusat-pusat cluster diperbarui dengan cara merata-rata posisi semua anggota cluster yang baru ditetapkan. Iterasi ini terus dilakukan sampai keanggotaan setiap cluster stabil atau tidak berubah lagi (Mishra dan Behera, 2012).

Proses clustering diawali dengan mengidentifikasi data yang akan diklaster x_{ij} ($i = 1, \dots, k; j = 1, \dots, m$) dan n adalah himpunan data yang akan dikeluarkan atau diklaster C_{kj} ($k = 1, \dots, k; j = 1, \dots, m$); maka jarak antara pusat setiap titik cluster dengan setiap datum, dapat digunakan rumus Euclidean untuk mengetahui jarak (Murti 2017).

Silhouette coefficient adalah salah satu teknik yang dilakukan untuk mengukur tingkat keoptimalan atau keabsahan sebuah *cluster* yang telah diperoleh dari *clustering* (Tanzil Furqon and Muflikhah 2016). *Silhouette Coefficient* akan memberikan hasil yang menunjukkan keakuratan suatu objek yang diteliti.

Tujuan penelitian ini adalah menganalisis metode *clustering k-means* pada zonasi daerah terdampak banjir di kota Medan dengan evaluasi *silhouette coefficient*.

2. METODE PENELITIAN

Jenis penelitian ini menggunakan pendekatan kuantitatif. Data kuantitatif adalah data yang diperoleh dari pengukuran numerik atau penghitungan yang dapat dianalisis secara statistik. Lokasi penelitian ini dilaksanakan di kota Medan yang terdiri dari 21 kecamatan.

Jenis data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder menggunakan data statistic, dokumen dan rekaman, literatur dan publikasi, dan database elektronik. Teknik analisis data diperoleh melalui manual *clustering* dengan *k-means*.

3. HASIL DAN PEMBAHASAN

Perhitungan Manual *Clustering* dengan K-Means

Adapun data yang digunakan diperoleh dari BPDB Provinsi Sumatera Utara atau diakses pada website www.dibi.bnppb.go.id berdasarkan tahun 2016-2023. Adapun objek data yang diambil adalah bencana banjir yang ada di seluruh 21 Kecamatan yang ada di kota Medan.

Setelah data diperoleh maka akan dibentuk *centroid* dengan jumlah *cluster* adalah 3, kelompok daerah aman dari banjir berada di *cluster* 1, kelompok daerah rawan dari banjir berada di *cluster* 2, dan kelompok daerah paling rawan dari banjir berada di *cluster* 3. Langkah ini disebut juga dengan langkah *clustering* K-Means. Selanjutnya, menentukan *centroid* yang diambil secara acak/random. Berikut ini adalah nilai *centroid* yang terbentuk:

Tabel 1. Nilai Centroid Awal

	Cluster		
	1	2	3
Terluka	0	56	97
Menderita	0	71	85
Mengungsi	0	86	94
Rumah Rusak Ringan	0	6	39

Rumah Rusak Sedang	1	29	57
Rumah Rusak Parah	0	0	78

Langkah selanjutnya yang akan dilakukan adalah dengan menentukan jarak masing-masing data terhadap *centroid* berdasarkan *Euclidean Distance*.

Iterasi 1:

1) Medan Tuntungan

Jarak *centroid* data ke 1 adalah:

$$\begin{aligned}
 d_{(1,1)} &= \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \\
 &= \sqrt{(76 - 0)^2 + (76 - 0)^2 + (80 - 0)^2 + (8 - 0)^2 + (59 - 1)^2 + (0 - 0)^2} \\
 &= \sqrt{(76)^2 + (76)^2 + (80)^2 + (8)^2 + (58)^2 + 0} \\
 &= \sqrt{5776 + 5776 + 6400 + 64 + 3364 + 0} \\
 &= \sqrt{21380} \\
 &= 146,22
 \end{aligned}$$

Jarak *centroid* data ke 2 adalah:

$$\begin{aligned}
 d_{(1,2)} &= \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \\
 &= \sqrt{(76 - 56)^2 + (76 - 71)^2 + (80 - 86)^2 + (8 - 6)^2 + (59 - 29)^2 + (0 - 0)^2} \\
 &= \sqrt{(20)^2 + (5)^2 + (-6)^2 + (2)^2 + (30)^2 + 0} \\
 &= \sqrt{400 + 25 + 36 + 4 + 900 + 0} \\
 &= \sqrt{1365} \\
 &= 36,946
 \end{aligned}$$

Perhitungan *Euclidean Distance* dilakukan pada seluruh sampel data dengan seluruh *centroid*. Tabel berikut menunjukkan hasil perhitungan *Euclidean Distance* seluruh sampel data dengan seluruh *centroid*.

Tabel 2. Perhitungan *Euclidean Distance* seluruh sampel data dengan seluruh *centroid*

Kecamatan	C1	C2	C3
Medan Tuntungan	146,22	36,946	88,131
Medan Johor	187,99	75,71	69,527
Medan Amplas	113	98,62	166,54
Medan Denai	112,08	95,216	164,09
Medan Area	24,021	113,57	179,97
Medan Kota	0	128,04	190,29
MedanMaimun	165,88	69,304	49,224
Medan Polonia	172,47	70,951	83,606
Medan Baru	190,29	99,489	0
Medan Selayang	175,84	90,515	21,61
Medan Sunggal	154,1	41,352	86,255

Kecamatan	C1	C2	C3
Medan Helvetia	165,02	50,547	86,366
Medan Petisah	21,024	117,94	182,19
Medan Barat	128,04	0	99,489
Medan Timur	182,23	74,639	87,378
Medan Perjuangan	12,042	121,09	184,74
Medan Tembung	35,114	105,88	174,07
Medan Deli	10,247	121,17	184,5
Medan Labuhan	35,128	100,95	170,55
Medan Marelan	148,43	39,912	87,264
Medan Belawan	197,26	99,975	99,514

Setelah menghitung *Euclidean Distance* ke masing-masing *centroid* pada iterasi 1, langkah selanjutnya adalah menentukan *cluster* mana yang paling dekat dengan setiap data. Kedekatan ini ditentukan berdasarkan jarak terpendek yang merupakan nilai terkecil dari jarak data tersebut ke C1, C2, dan C3.

Tabel 3. Jarak *Centroid*, Nilai Minimum, dan *Cluster* Iterasi 1

Kecamatan	C1	C2	C3	Minimum	Cluster
Medan Tuntungan	146,22	36,946	88,131	36,946	C2
Medan Johor	187,99	75,71	69,527	69,527	C3
Medan Amplas	113	98,62	166,54	98,62	C2
Medan Denai	112,08	95,216	164,09	95,216	C2
Medan Area	24,021	113,57	179,97	24,021	C1
Medan Kota	0	128,04	190,29	0	C1
Medan Maimun	165,88	69,304	49,224	49,224	C3
Medan Polonia	172,47	70,951	83,606	70,951	C2
Medan Baru	190,29	99,489	0	0	C3
Medan Selayang	175,84	90,515	21,61	21,61	C3
Medan Sunggal	154,1	41,352	86,255	41,352	C2
Medan Helvetia	165,02	50,547	86,366	50,547	C2
Medan Petisah	21,024	117,94	182,19	21,024	C1
Medan Barat	128,04	0	99,489	0	C2
Medan Timur	182,23	74,639	87,378	74,639	C2
Medan Perjuangan	12,042	121,09	184,74	12,042	C1
Medan Tembung	35,114	105,88	174,07	35,114	C1
Medan Deli	10,247	121,17	184,5	10,247	C1
Medan Labuhan	35,128	100,95	170,55	35,128	C1
Medan Marelan	148,43	39,912	87,264	39,912	C2
Medan Belawan	197,26	99,975	99,514	99,514	C3

Berdasarkan tabel 3 dapat dilihat bahwa nilai minimum pada kecamatan Medan Tuntungan adalah 36,946 yang berada pada *Cluster* ke-2, Sehingga kecamatan Medan Tuntungan digolongkan pada *cluster* ke-2. Hal tersebut juga berlaku pada seluruh kecamatan pada sampel acak. Penentuan *cluster* berdasarkan Perhitungan *Euclidean Distance* yang paling minimum.

Langkah selanjutnya dengan melanjutkan perhitungan iterasi ke-2 dengan membentuk *centroid* baru. *Centroid* terbaru diperoleh berdasarkan hasil perhitungan rata-rata setiap *cluster*, adapun hasil perhitungan dapat dilihat pada tabel berikut ini:

Adapun hasil *cluster* dari iterasi ke-2 adalah sebagai berikut:

Tabel 4. Hasil *Cluster* 1 dari iterasi ke-2

Kecamatan	Terluka	Menderita	Mengungsi	Rumah Rusak Ringan	Rumah Rusak Sedang	Rumah Rusak Parah
Medan Amplas	0	0	113	0	1	0
Medan Denai	0	4	112	0	2	0
Medan Area	0	0	24	0	0	0
Medan Kota	0	0	0	0	1	0
Medan Petisah	0	21	0	0	0	0
Medan Perjuangan	0	12	1	0	1	0
Medan Tembung	0	2	35	0	3	0
Medan Deli	2	10	1	0	1	0
Medan Labuhan	0	12	33	0	2	0
Rata-Rata	0,2222	6,7778	35,444	0	1,2222	0

Tabel 5. Hasil *Cluster* 2 dari iterasi ke-2

Kecamatan	Terluka	Menderita	Mengungsi	Rumah Rusak Ringan	Rumah Rusak Sedang	Rumah Rusak Parah
Medan Tuntungan	76	76	80	8	59	0
Medan Sunggal	81	87	80	9	57	0
Medan Helvetia	84	98	83	9	61	0
Medan Barat	56	71	86	6	29	0
Medan Marelan	88	65	86	8	52	0
Rata-Rata	77	79,4	83	8	51,6	0

Tabel 6. Hasil *Cluster* 3 dari iterasi ke-2

Kecamatan	Terluka	Menderita	Mengungsi	Rumah Rusak Ringan	Rumah Rusak Sedang	Rumah Rusak Parah
Medan Johor	91	110	92	37	71	15
Medan Maimun	102	87	75	11	42	46
Medan Polonia	85	91	80	40	80	0
Medan Baru	97	85	94	39	57	78
Medan Selayang	88	69	90	38	64	70
Medan Timur	120	80	90	9	66	0
Medan Belawan	144	76	82	7	76	0

Rata-Rata	103,86	85,429	86,143	25,857	65,143	29,857
------------------	--------	--------	--------	--------	--------	--------

Berdasarkan rata-rata dari tiap *cluster* yang didapat sehingga dibentuk *centroid* terbaru seperti pada tabel 7 ini:

Tabel 7. Nilai Centroid Terbaru

	Cluster		
	1	2	3
Terluka	0,2222	77	103,86
Menderita	6,7778	79,4	85,429
Mengungsi	35,444	83	86,143
Rumah Rusak Ringan	0	8	25,857
Rumah Rusak Sedang	1,2222	51,6	65,143
Rumah Rusak Parah	0	0	29,857

Berdasarkan data pada Tabel 7 dapat dilihat bahwa beberapa *cluster* masih ada yang berpindah dari iterasi 1 ke iterasi 2. Untuk itu, iterasi selanjutnya masih perlu dilakukan. Iterasi akan berhenti ketika tidak ada yang berpindah lagi antar *cluster* atau dengan kata lain *cluster* sudah menetap. Adapun hasil perhitungan iterasi ke-3 dapat dilihat pada tabel berikut ini:

Tabel 8. Centroid Iterasi Ke-3

Kecamatan	C1	C2	C3	Minimum	Cluster
Medan Tuntungan	126,18	8,7361	46,377	8,7361	C2
Medan Johor	168,99	51,534	34,389	34,389	C3
Medan Amplas	77,852	125,53	156,46	77,852	C1
Medan Denai	76,611	122,39	153,73	76,611	C1
Medan Area	13,358	135,8	166,58	13,358	C1
Medan Kota	36,088	147,47	176,58	36,088	C1
Medan Maimun	149,19	54,442	33,868	33,868	C3
Medan Polonia	155,15	45,146	41,673	41,673	C3
Medan Baru	172,13	87,33	51,631	51,631	C3
Medan Selayang	157,6	78,943	47,922	47,922	C3
Medan Sunggal	134,68	10,626	42,482	10,626	C2
Medan Helvetia	145,79	22,007	41,894	22,007	C2
Medan Petisah	38,211	137,68	167,82	38,211	C1
Medan Barat	102,95	32,176	71,35	32,176	C2
Medan Timur	164,2	45,899	38,486	38,486	C3
Medan Perjuangan	34,839	140,78	170,59	34,839	C1
Medan Tembung	5,122	129,03	160,58	5,122	C1
Medan Deli	34,641	140,68	170,26	34,641	C1
Medan Labuhan	5,8223	124,48	156,68	5,8223	C1
Medan Marelan	127,64	18,372	45,299	18,372	C2
Medan Belawan	182,41	71,4	55,517	55,517	C3

Berdasarkan tabel 8 dapat dilihat bahwa tidak ada lagi *cluster* yang berpindah ke *cluster* lain atau dengan kata lain *cluster* yang baru sama seperti *cluster* sebelumnya. Untuk itu, iterasi telah berakhir. Oleh karena itu, proses perhitungan diberhentikan pada iterasi ke-3. Hasil akhir dari pembentukan *cluster* pada iterasi ke-3 yaitu *Cluster* 1 memiliki 9 anggota data, *Cluster* 2 memiliki 5 anggota data, dan *Cluster* 3 memiliki 7 anggota data. Adapun hasil *cluster* dari iterasi ke-3 dapat dilihat pada tabel berikut:

Tabel 9. Pembentukan *Cluster* pada Iterasi 3

<i>Cluster 1</i>	Medan Amplas, Medan Denai, Medan Area, Medan Kota, Medan Petisah, Medan Perjuangan, Medan Tembung, Medan Deli, Medan Labuhan.
<i>Cluster 2</i>	Medan Tuntungan, Medan Sunggal, Medan Helvetia, Medan Barat, Medan Marelan.
<i>Cluster 3</i>	Medan Johor, Medan Maimun, Medan Polonia, Medan Baru, Medan Selayang, Medan Timur, Medan Belawan.

Berdasarkan pada Tabel 9 yang menunjukkan nilai pusat *cluster* (*centroid*) pada iterasi terakhir proses pembentukan *cluster* didapatkan sebagai berikut :

Cluster 1 : Daerah kecamatan yang aman dari banjir karena memiliki rata-rata jumlah korban yang sedikit. Terdiri dari 9 kecamatan yaitu kecamatan Medan Amplas, Medan Denai, Medan Area, Medan Kota, Medan Petisah, Medan Perjuangan, Medan Tembung, Medan Deli, dan Medan Labuhan

Cluster 2 : Daerah kecamatan yang rawan dari banjir karena memiliki rata-rata jumlah korban yang sedang. Terdiri dari 5 kecamatan yaitu kecamatan Medan Tuntungan, Medan Sunggal, Medan Helvetia, Medan Barat, dan Medan Marelan.

Cluster 3 : Daerah kecamatan yang paling rawan dari banjir karena memiliki rata-rata jumlah korban yang banyak. Terdiri dari 7 kecamatan yaitu kecamatan Medan Johor, Medan Maimun, Medan Polonia, Medan Baru, Medan Selayang, Medan Timur, dan Medan Belawan.

Contoh Perhitungan Manual Evaluasi *Clustering*

Untuk menganalisis hasil dari proses *clustering* menggunakan metode K-Means yang telah di dapat maka dibutuhkan perhitungan keakuratan hasil *cluster*. Untuk mengetahui keakuratan *cluster* dari masing-masing percobaan cluster di hitung nilai *Silhouette Coefficient*. Dalam penelitian ini , evaluasi yang digunakan pada hasil *Clustering* K-Means akan digunakan metode *Silhouette Coefficient*. Tujuan evaluasi ini dilakukan untuk mengetahui apakah *Clustering* yang dilakukan sudah akurat atau belum.

Silhouette Coefficient

Silhouette Coefficient adalah evalusai *clustering* yang digunakan untuk mengukur tingkat keakuratan hasil *clustering* data yang telah dilakukan. Adapun langkah pertama yang harus dilakukan adalah mencari nilai a_i atau menghitung nilai rata-rata perhitungan jarak data ke data yang lain dalam satu *cluster*.

Berikut perhitungan kecamatan Medan Tuntungan terhadap data lain pada *cluster* 2.

$$\begin{aligned}
 a_1 &= \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \\
 &= \frac{1}{4} \sqrt{(76 - 81)^2 + (76 - 87)^2 + (80 - 80)^2 + (8 - 9)^2 + (59 - 57)^2 + (0 - 0)^2} \\
 &= 3,072051
 \end{aligned}$$

Langkah selanjutnya yaitu menghitung nilai rata-rata jarak data ke- *i* dengan semua data pada *cluster* lain, dan pilihlah nilai minimumnya.

$$\begin{aligned}
 d(i, C) &= \frac{1}{|C|} \sum_{j \in C} d(i, j) \\
 &= \frac{1}{21} \sqrt{(76 - 81)^2 + (76 - 87)^2 + (80 - 80)^2 + (8 - 9)^2 + (59 - 57)^2 + (0 - 0)^2} \\
 &= 0,585153
 \end{aligned}$$

Perhitungan $d(i, C)$ akan dilanjutkan pada data seluruh data pada *cluster* yang lain. Langkah selanjutnya yaitu menghitung nilai $b(i)$ seperti berikut ini.

$$\begin{aligned}
 b_1 &= \min_{C \neq A} d(i, C) \\
 &= \min(0,585153; 1,128883; 1,759329; 0,890871) \\
 &= 0,585153
 \end{aligned}$$

Setelah nilai minimum telah diperoleh maka langkah selanjutnya adalah menghitung nilai *Silhouette Coefficient*.

$$\begin{aligned}
 s_1 &= \frac{b(i) - a(i)}{\max(a(i), b(i))} \\
 &= \frac{0,585153 - 5,728059}{\max(5,728059; 0,585153)} \\
 &= 0,897844
 \end{aligned}$$

Adapun hasil perhitungan pada seluruh data dapat dilihat pada tabel berikut ini :

Tabel 10. Hasil Perhitungan *Silhouette Coefficient* seluruh data

Kecamatan	a_i	b_i	s_i
Medan Tuntungan	5,728059	0,585153	0,897844
Medan Johor	9,756354	1,393213	0,857199
Medan Amplas	11,0194	0,202031	0,981666
Medan Denai	10,88509	0,202031	0,98144
Medan Area	4,857499	0,55123	0,88652
Medan Kota	5,666429	0,48795	0,913888
Medan Maimun	9,901583	2,343999	0,76327
Medan Polonia	10,49734	1,393213	0,867279
Medan Baru	11,41277	1,029056	0,909833
Medan Selayang	11,00165	1,029056	0,906463
Medan Sunggal	5,65518	0,585153	0,896528
Medan Helvetia	7,580417	0,592852	0,921792
Medan Petisah	5,870857	0,43383	0,926104
Medan Barat	10,54734	1,759329	0,833197

Kecamatan	a_i	b_i	s_i
Medan Timur	9,974495	1,312767	0,868388
Medan Perjuangan	5,33418	0,134687	0,97475
Medan Tembung	5,038977	0,48795	0,903165
Medan Deli	5,318051	0,134687	0,974674
Medan Labuhan	4,988014	0,48795	0,902175
Medan Marelan	7,349784	0,890871	0,878789
Medan Belawan	12,14601	1,312767	0,891918
	<i>Silhouette Coefficient</i>		0,901756

Hasil akhir *Silhouette Coefficient* adalah dengan menghitung nilai rata-ratanya. Adapun hasil perhitungan rata-ratanya adalah $0,901756 \approx 0,9$. Dalam menentukan tingkat keakuratan nilai *Silhouette Coefficient* jika nilainya ≥ 0 maka hasilnya dinyatakan cukup signifikan. Karena nilai *Silhouette Coefficient* yang diperoleh ≥ 0 maka hasil *clustering* dengan K-Means dalam penelitian ini dapat dinyatakan dengan cukup signifikan.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan diperoleh bahwa Cluster 1 : Cluster ini berisi kecamatan yang memiliki tingkat banjir yang aman karena rata-rata jumlah korban yang sedikit. Terdiri dari 9 kecamatan yaitu kecamatan Medan Amplas, Medan Denai, Medan Area, Medan Kota, Medan Petisah, Medan Perjuangan, Medan Tembung, Medan Deli, dan Medan Labuhan. Cluster 2 : Cluster ini berisi kecamatan yang memiliki tingkat banjir yang rawan karena rata-rata jumlah korban yang sedang. Terdiri dari 5 kecamatan yaitu kecamatan Medan Tuntungan, Medan Sunggal, Medan Helvetia, Medan Barat, dan Medan Marelan. Cluster 3 : Cluster ini berisi kecamatan yang memiliki tingkat banjir yang paling rawan karena nilai rata-rata jumlah korban yang tinggi. Terdiri dari 7 kecamatan yaitu kecamatan Medan Johor, Medan Maimun, Medan Polonia, Medan Baru, Medan Selayang, Medan Timur, dan Medan Belawan. Sehingga diperoleh 3 iterasi dengan keakuratan nilai *Silhouette Coefficient* yaitu 0,9. Karena nilai *Silhouette Coefficient* yang diperoleh ≥ 0 maka hasil *clustering* dengan K-Means dalam penelitian ini dapat dinyatakan dengan cukup signifikan.

DAFTAR REFERENSI

- Afifah, N., Rini, D. C., & Lubab, A. (2016). Pengklasteran lahan sawah di Indonesia menggunakan Fuzzy C-Means. *Journal Matematika MANTIK*, 02(01), 40–45.
- Alamsyah, B., Nuraini, C., Kuswandi, & Suwarno, B. (2018). *Strategi manajemen mitigasi bencana pesisir pantai timur Sumatera Utara*. Medan: UNPRIPRESS.
- Aminah, S. (2018). Implementasi Permendikbud No 17 Tahun 2017 tentang penerimaan peserta didik baru di SMA Negeri 1 Prambon Nganjuk Tahun 2017/2018. (Undergraduate thesis, IAIN Kediri). Retrieved February 19, 2022, from <http://etheses.iainkediri.ac.id/766/>
- Darwiyanto, E., Binawan, P. B., & Junaedi, D. (2017). Aplikasi GIS klasifikasi tingkat kerawanan banjir wilayah Kabupaten Bandung menggunakan metode Weighted Product. *Journal on Computing*, 2(1), 59–70.
- Dehotman, B. J. (2018). Peningkatan hasil evaluasi clustering Davies-Bouldin index dengan penentuan titik pusat cluster awal algoritma K-Means.
- Dewi, A. I. D., & Pramita, A. K. D. (2019). Analisis perbandingan metode Elbow dan Silhouette pada algoritma clustering K-Medoids dalam pengelompokan produksi kerajinan Bali. *Journal Matrix*, 9(3), 102–109.
- Everitt, B. S., Sabine, L., Morven, L., & Daniel, S. (2011). *Cluster analysis* (5th ed.).
- Handoyo, R. R., & Nasution, S. N. (2014). Perbandingan metode clustering menggunakan metode single linkage dan K-Means pada pengelompokan dokumen. *JSM STMIK Mikroskil*, 15(2), 73–82.
- Hardaningrum, O., Oxtavi, S., & Supriyana, E. (2016). Zonasi rawan bencana gempa bumi Kota Malang berdasarkan analisis Horizontal Vertical to Spectral Ratio (HVSR). 27–28. Retrieved from https://www.researchgate.net/profile/AdeHeryana/publication/338537206_Pengertian_dan_Jenis_Bencana/links/5e1a694c4585159aa4c8bbe4/Pengertian-dan-Jenis-Bencana.pdf
- Karputri, L. N., & Yustanti, W. (2022). Analisis klatering buku sebagai evaluasi untuk peningkatan minat baca perpustakaan SMAN 1 Grogol. *Journal of Emerging Information System and Business Intelligence (JEISBI)*, 3(3), 94–101.
- Sari, N. N. (2018). 15 Universitas Islam Indonesia K-Affinity Propagation (K-Ap) clustering untuk klasifikasi gempa bumi. Yogyakarta.
- Tampubolon, K. (2018). Aplikasi sistem informasi geografis (SIG) sebagai penentuan kawasan rawan banjir di Kota Medan. *Journal Pembangunan Perkotaan*, 6(2), 63–68.