

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KLASTERISASI DATA GEMPA DI INDONESIA MENGUNAKAN K-MEANS CLUSTERING

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

HAYYUDI GIANNANDA ASFA

NIM. 11850112515



UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU**

2025



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**KLASTERISASI DATA GEMPA DI INDONESIA
MENGUNAKAN K-MEANS CLUSTERING**

TUGAS AKHIR

Oleh

HAYYUDI GIANNANDA ASFA

NIM. 11850112515

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 30 Juni 2025

Pembimbing I,

Pembimbing II,

Novriyanto, S.T., M.Sc.
NIP. 19771128 200710 1 003

Dr. Elin Haerani, S.T., M.Kom.
NIP. 19810323 200710 2 003



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

**KLASTERISASI DATA GEMPA DI INDONESIA
MENGUNAKAN K-MEANS CLUSTERING**

Oleh

HAYYUDI GIANNANDA ASFA

NIM. 11850112515

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 30 Juni 2025

Mengesahkan,

Ketua Jurusan,

Dekan,

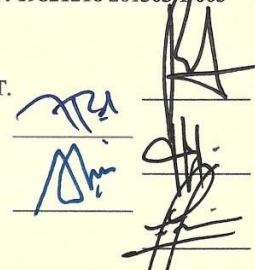
Dr. Hartono, M.Pd.
NIP. 19640301 199203 1 003


Iwan Iskandar, S.T., M.T.

NIP. 19821216 201503 1 003

DEWAN PENGUJI

Ketua : Dr. Rahmad Abdillah, S.T., M.T.
Pembimbing I : Novriyanto, S.T., M.Sc.
Pembimbing II : Dr. Elin Haerani, S.T., M.Kom.
Penguji I : Dr. Alwis Nazir, M.Kom.
Penguji II : Fitri Insani, S.T., M.Kom.





Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lampiran Surat :
 Nomor : Nomor 25/2021
 Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : HAYYUDI GIANNANDA ASFA
 NIM : 11850112515
 Tempat/Tgl. Lahir : RENGAT / 20 NOVEMBER 2000
 Fakultas/Pascasarjana : SAINS DAN TEKNOLOGI
 Prodi : TEKNIK INFORMATIKA

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:

KLASTERISASI DATA GEMPA DI INDONESIA MENGGUNAKAN
 K-MEANS CLUSTERING

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 12 JULI 2025
 membuat pernyataan



HAYYUDI GIANNANDA ASFA
 NIM : 11850112515

*pilih salah satu sesuai jenis karya tulis



LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul “KLAUSTERISASI DATA GEMPA DI INDONESIA MENGGUNAKAN K-MEANS CLUSTERING” adalah gagasan asli dari saya sendiri dan belum pernah dijadikan Tugas Akhir atau sejenisnya di Universitas Islam Negeri Sultan Syarif Kasim Riau maupun di perguruan tinggi lain.
2. Dalam Tugas Akhir ini TIDAK terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali tertulis dengan jelas dan dicantumkan sebagai referensi di dalam Daftar Pustaka.
3. Dalam Tugas Akhir ini TIDAK terdapat penggunaan Kecerdasan Buatan Generatif (Generative AI) yang bertentangan dengan ketentuan dan peraturan yang berlaku.
4. Saya bersedia menerima sanksi sesuai ketentuan yang berlaku apabila di kemudian hari terbukti bahwa Tugas Akhir ini melanggar kode etik maupun peraturan yang berlaku, termasuk plagiat ataupun pelanggaran hak cipta.

Demikianlah pernyataan ini dibuat dengan sebenarnya.

Pekanbaru, 30 Juni 2025

Yang membuat pernyataan,

HAYYUDI GIANNANDA ASFA

11850112515



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

Penulis mempersembahkan tulisan ini kepada kedua orangtua penulis yang membesarkan penulis dengan penuh kasih sayang dan tidak hentinya memberikan doa kepada penulis dalam segala hal yang dilakukan penulis. Semua yang dilakukan penulis tidak lepas dari doa dan dukungan dari kedua orangtua sehingga penulis menjadi pribadi saat ini. Tentunya tak terhitung jasa mereka dalam kehidupan penulis. Penulis mengucapkan terimakasih. Selain itu, kepada kedua saudara penulis yang juga memberikan banyak doa dan dukungan kepada penulis, penulis ucapkan terimakasih.

Penulis mengucapkan terimakasih kepada dosen pembimbing tugas akhir penulis, yang selalu membimbing penulis sehingga laporan ini dapat terselesaikan. Begitu banyak bimbingan, arahan, dan masukan serta bantuan-bantuan yang diberikan selama pelaksanaan tugas akhir penulis, sehingga penulis bisa menyelesaikan dengan baik.

Kepada teman-teman penulis yang banyak menemani dan membantu penulis selama perkuliahan hingga pelaksanaan tugas akhir, penulis juga mengucapkan banyak terimakasih sehingga penulis dapat lancar dalam perkuliahan hingga tugas akhir.

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ABSTRAK

Indonesia merupakan wilayah yang sering mengalami gempa seismik karena dikelilingi oleh empat lempeng utama dunia. Gempa dapat membuat kerusakan pada daerah yang mengalaminya bahkan membahayakan jiwa manusia. Walaupun saat ini teknologi sudah semakin maju, tetapi gempa bumi belum dapat diprediksi waktu dan tempat terjadinya. Meskipun begitu, teknologi yang sudah ada saat ini mampu memetakan daerah rawan gempa. Maka dilakukan pendekatan dengan menggunakan penghitungan *clustering*. Pada penelitian ini digunakan algoritma *K-Means Clustering* untuk mengolah data. Data berisi 92887 data gempa bumi yang terjadi di Indonesia. Penghitungan *K-Means* menggunakan *Euclidean Distance* dan *Manhattan Distance*. Pada *Euclidean Distance*, didapatkan hasil *cluster* 1 memiliki 49838 *record*, *cluster* 2 berisi 33126 *record*, dan *cluster* 3 dengan 7188 *record*. Sedangkan melalui penghitungan *Manhattan Distance*, didapatkan hasil *cluster* 1 memiliki 36525 *record*, *cluster* 2 dengan 46676 *record*, dan *cluster* 3 berisikan 6951 *record*. Didapatkan *cluster* berjumlah 3 melalui penghitungan *Elbow Method*. Nilai *Silhouette Score* pada penghitungan *Euclidean Distance* penelitian ini adalah 0,469702780513585 dan pada *Manhattan Distance* nilai *Silhouette Score* yaitu 0,44916422524456295. Dapat disimpulkan bahwa nilai *Silhouette Score* penghitungan *Euclidean Distance* menghasilkan kualitas *clustering* yang lebih baik dari penghitungan *Manhattan Distance* pada penelitian klasterisasi data gempa di Indonesia menggunakan *K-Means Clustering* ini.

Kata kunci: euclidean, gempa, k-means, klasterisasi, manhattan

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ABSTRACT

Indonesia is an area that often experiences seismic earthquakes because it is surrounded by four major plates of the world. Earthquakes can cause damage to areas that experience them and even endanger human lives. Although technology is currently more advanced, earthquakes cannot be predicted when and where they occur. Even so, existing technology is able to map earthquake-prone areas. So an approach is carried out using clustering calculations. In this study, the K-Means Clustering algorithm was used to process the data. The data contains 92887 earthquake data that occurred in Indonesia. The K-Means calculation uses Euclidean Distance and Manhattan Distance. In Euclidean Distance, cluster 1 has 49838 records, cluster 2 contains 33126 records, and cluster 3 with 7188 records. In Manhattan Distance, cluster 1 has 36525 records, cluster 2 with 46676 records, and cluster 3 contains 6951 records. There is 3 cluster used after the Elbow Method calculation. The Silhouette Score in the Euclidean Distance is 0,469702780513585 and in the Manhattan Distance the Silhouette Score is 0,44916422524456295. It concluded that the Silhouette Score of the Euclidean Distance calculation produces better clustering quality than the Manhattan Distance calculation in this study of earthquake data clustering in Indonesia using K-Means Clustering.

Keywords: euclidean, earthquake, k-means, clustering, manhattan

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Alhamdulillah penulis ucapkan kehadiran Allah *Subhanahu wa ta'ala*, karena berkat taufik serta hidayah-Nya Tugas Akhir ini dapat diselesaikan. Tak lupa pula untuk mengucapkan shalawat kepada Nabi Muhammad *Sholallohu 'alaihi wa salam*, semoga kita mendapat syafat di akhirat kelak.

Ada banyak yang membantu penulis dalam penyusunan laporan ini dengan memberi banyak motivasi sehingga laporan ini bisa diselesaikan. Terdapat banyak pihak yang membantu penulis dan tidak bisa untuk penulis menyebutkannya satu-persatu, tetapi pada kesempatan ini penulis hanya dapat mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Hj. Leny Nofianti, M.S., S.E., M.Si., Ak, CA., Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd., Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Iwan Iskandar, S.T., M.T., Kepala Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Dr. Elin Haerani, S.T., M.Kom., Penasihat Akademik sekaligus pembimbing 2 tugas akhir penulis.
5. Bapak Novriyanto, S.T., M.Sc., pembimbing 1 tugas akhir penulis
6. Bapak Dr. Alwis Nazir, M.Kom. dan Ibu Fitri Insani, S.T., M.Kom., penguji tugas akhir penulis
7. Orang tua penulis, Bapak Abdul Fatah dan Ibu Asmayanti yang selalu memberikan do'a, dukungan, serta arahan kepada penulis.
8. Semua yang banyak membantu dan mendukung penulis sehingga bisa menjadi seperti sekarang ini.

Penulis sadar bahwa masih memiliki berbagai kekurangan dan kemungkinan terdapat sejumlah kesalahan dalam laporan ini. Oleh karena itu, penulis sangat



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

mengharapkan masukan dan saran yang membangun sebagai bahan evaluasi dan pembelajaran di masa mendatang. Penulis juga berharap bahwa laporan ini bisa memberikan kegunaan bagi setiap pembaca.

Pekanbaru, 30 Juni 2025

Penulis



UIN SUSKA RIAU

DAFTAR ISI

LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN.....	iii
SURAT PERNYATAAN KEASLIAN	iv
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	v
LEMBAR PERNYATAAN	vi
LEMBAR PERSEMBAHAN.....	vii
ABSTRAK	viii
ABSTRACT	ix
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR RUMUS	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
BAB 2 KAJIAN PUSTAKA	5
2.1 Gempa Bumi.....	5
2.2 Klasterisasi	5
2.3 K-Means Clustering	6

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

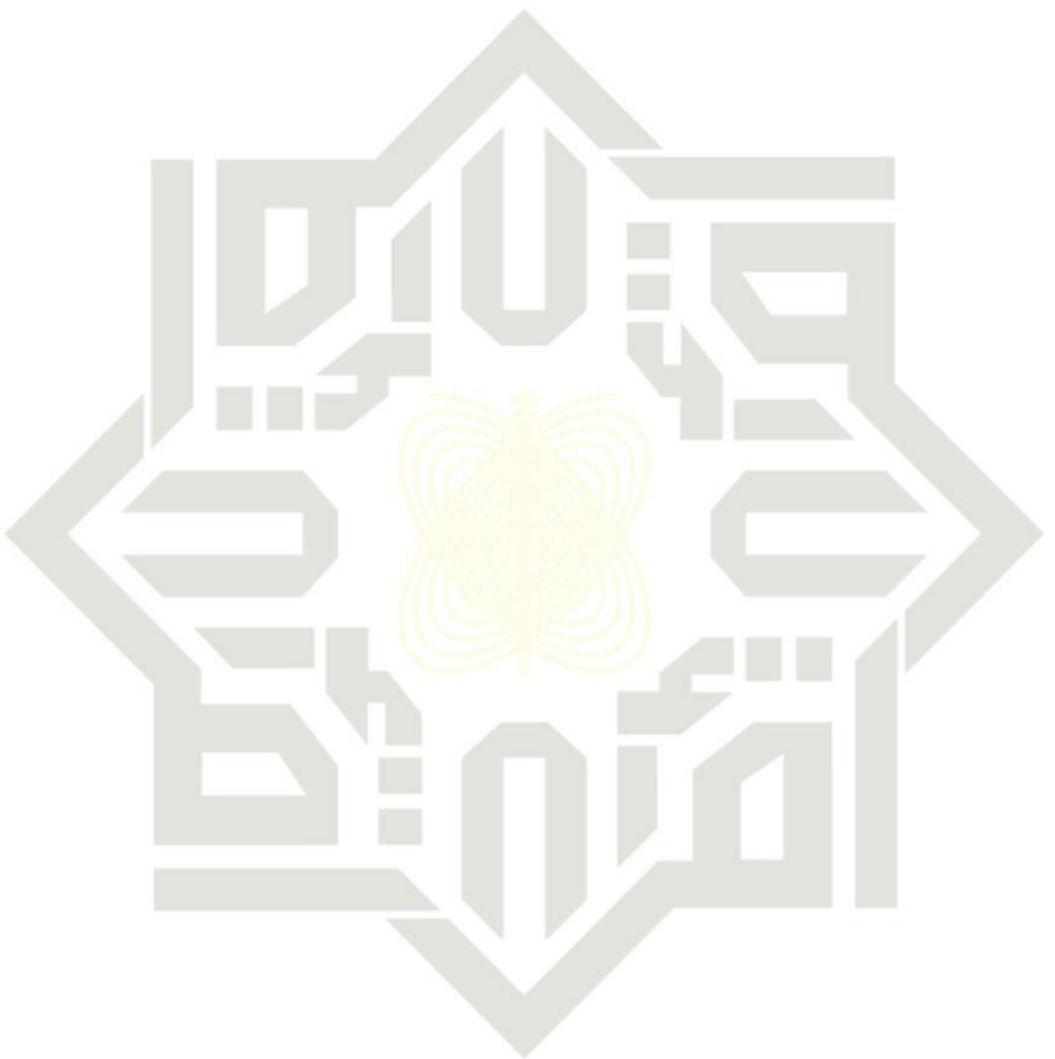
2.4	Euclidean Distance	7
2.5	Manhattan Distance	7
2.6	Knowledge Discovery in Database (KDD)	7
2.7	Elbow Method	8
2.8	Silhouette Score	9
2.9	Penelitian Terkait	9
BAB 3 METODOLOGI PENELITIAN		12
3.1	Tahapan Penelitian	12
3.2	Identifikasi Masalah	12
3.3	Studi Pustaka	13
3.4	Perumusan Masalah	13
3.5	Pengumpulan Data	13
3.6	Analisis Data	13
3.6.1	Analisis Proses KDD	14
3.7	Kesimpulan dan Saran	15
BAB 4 PEMBAHASAN		16
4.1	Data Penelitian	16
4.2	Data Selection	16
4.3	Data Preprocessing	17
4.4	Transformation	18
4.5	Data Mining	18
4.6	Evaluation	27
BAB 5 PENUTUP		29
5.1	Kesimpulan	29
5.2	Saran	30



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA	31
DAFTAR RIWAYAT HIDUP	34



UIN SUSKA RIAU

DAFTAR GAMBAR

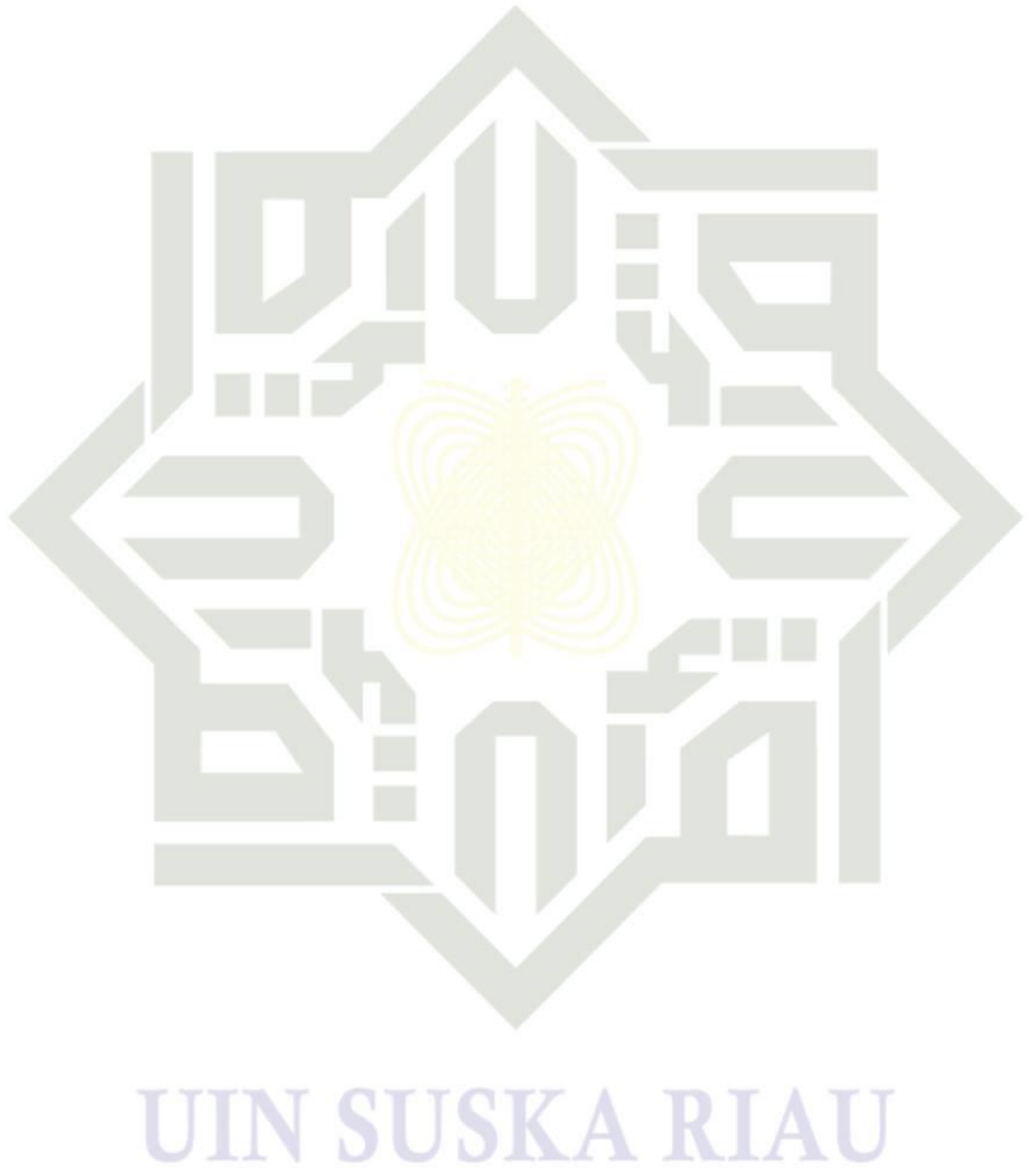
Gambar 4. 1 Data Selection.....	17
Gambar 4. 2 Data Setelah Preprocessing	17
Gambar 4. 3 Transformation Data	18
Gambar 4. 4 Visualisasi Data K-Means Euclidean Distance	19
Gambar 4. 5 Cluster 1 Euclidean Distance	20
Gambar 4. 6 Cluster 2 Euclidean Distance	21
Gambar 4. 7 Cluster 3 Euclidean Distance	22
Gambar 4. 8 Visualisasi Data K-Means Manhattan Distance	23
Gambar 4. 9 Cluster 1 Manhattan Distance	24
Gambar 4. 10 Cluster 2 Manhattan Distance	25
Gambar 4. 11 Cluster 3 Manhattan Distance	26
Gambar 4. 12 Elbow Method	27

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR TABEL

Tabel 2. 1 Penelitian terkait.....	9
Tabel 4. 1 Data Penelitian	16



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



DAFTAR RUMUS

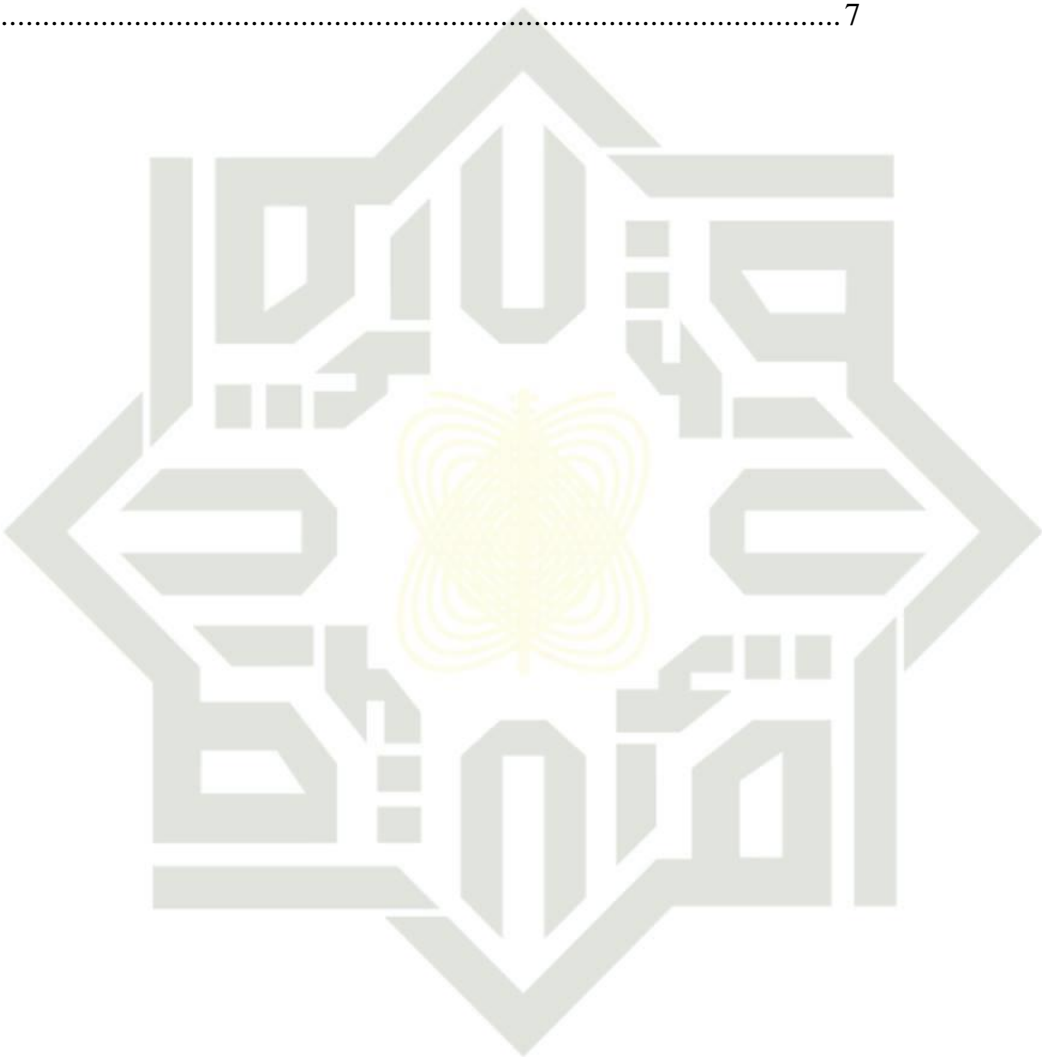
© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

.....	6
.....	7
.....	7



UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang dikenal memiliki alam yang indah. Walaupun begitu, alam yang indah ini juga memiliki kemungkinan-kemungkinan terjadinya bencana alam. Hal ini tak lepas dari status Indonesia yang terletak di “*ring of fire*” pasifik. Istilah tersebut untuk menunjukkan bahwa Indonesia berada pada daerah pertemuan lempeng tektonik. Sehingga Indonesia merupakan daerah pertemuan dari rangkaian jalur pegunungan dunia yang dikenal dengan Sirkum Mediterania dan Sirkum Pasifik. Daerah Cincin Api Pasifik merupakan zona yang mengelilingi cekungan Samudra Pasifik dan dikenal dengan karakteristik aktivitas gempa bumi serta letusan gunung berapi yang terjadi secara intens di sepanjang daerah tersebut. Dengan kenyataan bahwa alam Indonesia berada di daerah “*ring of fire*” ini, menjadikan wilayah Indonesia sering dilanda gempa bumi tektonik dan memiliki banyak gunung api (Oktavian & Adiarto, 2017).

Kejadian bencana alam disebabkan oleh berbagai aktivitas objek alam yang bersifat destruktif di muka bumi. Dampak kejadian ini menyebabkan berbagai masalah seperti terganggunya aktivitas manusia. Selain itu, sebagian besar bencana alam menyebabkan rusaknya objek vital di daerah sekitarnya, seperti tempat tinggal, fasilitas umum, tempat kerja, dan lain-lain. Kejadian bencana alam juga berpeluang besar menyebabkan jatuhnya korban jiwa termasuk luka-luka hingga meninggal dunia. Bahaya bencana alam juga diperparah dengan waktu kejadiannya yang tak terduga atau datang secara tiba-tiba (I. N. Setiawan, Krismawati, Pramana, & Tanur, 2022).

Di Indonesia, gempa bumi merupakan bencana alam yang umum terjadi, disebabkan oleh aktivitas tektonik berupa pergerakan lempeng-lempeng bumi yang menimbulkan getaran dan menghasilkan gelombang seismik di permukaan (Tania,



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Handhayani, & Hendryli, 2023). Karena posisi geografis Indonesia yang berposisi di pertemuan empat lempeng tektonik utama, yakni Pasifik, Indo-Australia, Eurasia, dan Samudera Filipina, Indonesia termasuk area yang kerap merasakan gempa seismik (Saleh, Mandar, & Noh, 2023). Gempa bumi dapat menimbulkan kerusakan parah pada berbagai hal, termasuk dapat mengakibatkan munculnya banyak korban jiwa dan cedera. Gempa bumi difaktori beberapa hal seperti magnitudo gempa, kedalaman pusat gempa serta jarak antara lokasi terjadinya gempa dengan permukaan tanah (Duha, Laia, Huda, & Jasuma, 2023). Walaupun saat ini teknologi sudah semakin maju, tetapi gempa bumi belum dapat diprediksi waktu dan tempat terjadinya. Meskipun begitu, teknologi yang sudah ada saat ini mampu memetakan daerah rawan gempa dan merancang bangunan tahan gempa (Qothrunnada, Utami, & Rizky, 2022). Maka dilakukan pendekatan dengan menggunakan penghitungan *clustering*.

Clustering adalah teknik pengelompokan data yang digunakan dalam mengorganisasi, observasi, atau entitas ke dalam beberapa kelompok atau kelas. *Clustering* dapat digunakan ke dataset tanpa kategori tertentu, karena metode ini bekerja tanpa memerlukan data berlabel (Pambudhi, Homaidi, & Santoso, 2024). *Clustering* menyediakan beragam pendekatan yang bisa digunakan sesuai dengan kebutuhan analisis. Teknik pengelompokan data yang populer dan sering diterapkan dalam berbagai analisis satu diantaranya yaitu *K-Means Clustering* (Heraldi, Aprilia, & Pratiwi, 2019).

K-Means Clustering merupakan bagian dari *Unsupervised Learning* karena sistem partisi digunakan dalam proses pembagian data ke dalam kelompok-kelompok tertentu. *Unsupervised Learning* merupakan algoritma data mining mencari informasi pada suatu atribut. Data akan dibagi menjadi beberapa *cluster* dengan melakukan pengelompokan berbasis jarak yaitu *K-Means Clustering* (Musid, Dikananda, & Fathurrohman, 2023). Algoritma *K-Means* merupakan metode *clustering non-hierarchical* dengan kecepatan penghitungan yang relatif cepat. *K-Means Clustering* adalah *clustering* yang membutuhkan k parameter masukan dan mempartisi sekumpulan n objek ke dalam k klaster sedemikian rupa



Hak Cipta Ditanggung Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sehingga kesamaan anggota kluster tinggi dan kesamaan dengan anggota kluster lainnya rendah. Tingkat kesamaan antara anggota dan *cluster* ditentukan berdasarkan seberapa dekat data dengan dengan rata-rata pusat kluster (Widyawati & Fauzy, 2024).

Berdasarkan penjelasan di atas maka penelitian akan menggunakan algoritma *K-Means Clustering* untuk menerapkan klasterisasi data gempa di Indonesia. Penelitian yang dilakukan diharapkan dapat membantu memberikan informasi mengenai daerah rawan gempa dan dapat memberikan kontribusi dalam membantu pemetaan daerah rawan gempa serta usaha dalam mitigasi bencana gempa bumi di Indonesia.

1.2 Rumusan Masalah

Dari masalah yang terdapat di latar belakang di atas, dapat dirumuskan permasalahan dalam penelitian ini adalah bagaimana menerapkan algoritma *K-Means* untuk klasterisasi data gempa di Indonesia.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Data yang diolah adalah data gempa yang terjadi di Indonesia.
2. Data berasal dari situs Kaggle.
3. Penghitungan dalam klasterisasi penelitian ini adalah algoritma *K-Means Clustering*.
4. Penghitungan *K-Means Clustering* menggunakan *Euclidean Distance* dan *Manhattan Distance*.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menerapkan algoritma *K-Means* dalam klasterisasi data gempa di Indonesia.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.5 Manfaat Penelitian

Manfaat penelitian merupakan penjelasan nilai serta kontribusi dari penelitian yang dilakukan. Manfaat penelitian ini adalah:

1. Hasil klasterisasi dari penelitian ini diharapkan dapat memberikan informasi mengenai daerah-daerah yang termasuk rawan hingga tidak rawan akan terjadinya bencana gempa bumi.
2. Penelitian ini diharapkan dapat meningkatkan waspada kepada masyarakat pada daerah rawan bencana gempa bumi.
3. Penelitian ini diharapkan dapat membantu pihak terkait akan bencana alam untuk perencanaan mitigasi bencana gempa bumi dan edukasi masyarakat yang berada pada daerah rawan gempa bumi.



BAB 2

KAJIAN PUSTAKA

2.1 Gempa Bumi

Gempa bumi adalah fenomena getaran pada lapisan litosfer dan permukaan bumi yang dikarenakan berbagai hal, diantaranya tabrakan antar lempeng tektonik, aktivitas patahan aktif, letusan gunung berapi, maupun runtuh batuan dalam volume yang begitu besar. Umumnya, kejadian gempa bumi biasanya muncul dikarenakan adanya pergerakan kerak bumi (lempeng bumi). Alat ukur yang digunakan untuk mengukur gempa bumi bernama Seismometer atau juga dikenal sebagai Seismograf. Sedangkan untuk mengukur kekuatan sebuah kejadian gempa bumi diukur dengan skala yang paling umum digunakan untuk seluruh dunia bernama *moment magnitude scale* (Ruyani, 2023).

2.2 Klasterisasi

Clustering (klasterisasi) adalah metode analisis data yang berfungsi mengelompokkan data ke kelompok tertentu berdasarkan kemiripan maupun kedekatan karakteristiknya. Dalam sebuah klaster, data di dalamnya memiliki tingkat kemiripan yang tinggi pada satu klaster dan perbedaan yang jelas dibandingkan dengan data di klaster lainnya. Tujuan utama klastering adalah untuk menemukan pola baru dalam data yang tidak memiliki label. Metode ini disebut sebagai *unsupervised learning* karena tidak memerlukan data label untuk melakukan pengelompokan (Fajri et al., 2024).

Clustering adalah proses membagi sekumpulan objek data ke beberapa kelompok yang dikenal dengan istilah *cluster*. Data yang tergolong dalam satu klaster menunjukkan kemiripan antar elemen, serta perbedaan yang jelas dari klaster lainnya. *Clustering* sering disebut juga sebagai segmentasi data karena proses ini membagi kumpulan data ke dalam beberapa grup sesuai kesamaan antar data (Z. Setiawan et al., 2023).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.3 K-Means Clustering

K-Means adalah metode *clustering* yang banyak digunakan dalam *data mining* dan *machine learning*. Metode ini membagi data ke dalam K kluster berdasarkan kedekatan jarak antara data dengan pusat kluster. *K-Means* bekerja dengan cara menentukan K pusat kluster (*centroids*) secara acak dan menempatkan data dalam kelompok yang sama berdasarkan perhitungan jarak minimum dengan pusat kluster tersebut. Langkah ini dilakukan berulang kali sampai posisi pusat kluster stabil dan tidak mengalami perubahan yang berarti (Fajri et al., 2024).

Tahap *clustering* pada *K-Means* adalah sebagai berikut (Dosbing, 2024):

1. Menetapkan nilai k sebagai acuan jumlah kluster yang akan dibentuk.
2. Mengambil nilai acak sebagai titik pusat awal (*centroid*) sebanyak jumlah kluster yang ditentukan.
3. Menghitung jarak setiap data terhadap seluruh *centroid* menggunakan rumus pengukuran jarak, lalu menentukan titik terdekat untuk tiap data.
4. Mengelompokkan data ke dalam kluster berdasarkan kedekatannya dengan *centroid* (jarak terkecil).
5. Melakukan pembaruan posisi *centroid* berdasarkan rata-rata nilai dari data dalam masing-masing kluster, dengan menggunakan rumus:

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j \quad 2.1$$

$\mu_j(t+1)$: *centroid* baru pada iterasi ke- $(t+1)$, N_{sj} : banyak data pada *cluster* S_j

6. Mengulangi proses mulai dari langkah ke-3 hingga langkah ke-5 secara berulang, sampai tidak ada lagi perubahan anggota dalam masing-masing kluster.

Tujuan utama dari proses *clustering* data adalah untuk meminimalkan fungsi objektif yang ditetapkan selama pengelompokan, yang umumnya bertujuan untuk

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

mengurangi variasi di dalam setiap kluster sekaligus meningkatkan perbedaan antar kluster (Nurhayati, 2022).

2.4 Euclidean Distance

Euclidean Distance adalah cara menghitung jarak antara dua titik dalam suatu ruang, dengan mempertimbangkan posisi titik-titik tersebut dalam dimensi yang sama (Sutirta & Noviandi, 2024). Rumus yang digunakan adalah (Dosbing, 2024):

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \quad 2.2$$

2.5 Manhattan Distance

Manhattan Distance adalah metode perhitungan jarak yang digunakan untuk mengukur jarak absolut antara dua titik koordinat suatu objek. Penghitungan jaraknya didasarkan pada selisih nilai antara dua titik koordinat yang dibandingkan. Persamaan yang digunakan dalam penghitungan ini yaitu (Sutirta & Noviandi, 2024) :

$$d = \sum_{i=1}^n |x_i - y_i| \quad 2.3$$

2.6 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan proses atau pendekatan untuk mengekstraksi informasi atau pengetahuan yang berguna dari *database* yang telah ada. Adapun langkah-langkah dalam proses KDD adalah sebagai berikut (Alghifari & Juardi, 2021).



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Data Selection

Data Selection adalah proses pemilihan data dari sebuah dataset yang nantinya akan dilakukan pengolahan.

2. Data Preprocessing

Data Preprocessing adalah tahap awal dalam proses pengolahan data yang bertujuan untuk membersihkan data dari gangguan seperti noise maupun ketidakkonsistenan, sehingga data yang digunakan menjadi lebih akurat dan siap untuk dianalisis lebih lanjut.

3. Transformation

Transformation merupakan tahap di mana data diubah atau disesuaikan agar cocok dengan model atau algoritma yang akan digunakan dalam proses analisis data selanjutnya.

4. Data Mining

Tahapan ini merupakan proses eksplorasi dan ekstraksi pengetahuan untuk membentuk sebuah model yang mampu menghasilkan informasi yang bernilai dan bermanfaat.

5. Evaluation

Pada *Evaluation*, dilakukan evaluasi yang merepresentasikan hasil model yang telah didapatkan.

2.7 Elbow Method

Elbow Method atau Metode *Elbow* merupakan sebuah teknik yang membandingkan hasil antara jumlah *cluster* yang membentuk *elbow* (siku) pada sebuah titik yang digunakan untuk menentukan nilai *cluster* yang paling optimal dalam pengolahan data (Safira & Castaka, 2024). *Elbow Method* ini dihitung berdasarkan kuadrat jarak antar titik sampel di pusat *cluster* dan pusat *cluster* untuk memperoleh serangkaian nilai *k* yang ditentukan dengan menetapkan kurva yang melengkung signifikan ke bawah, di titik tersebut menunjukkan nilai *k* yang paling optimal (Sutirta & Noviandi, 2024).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.8 Silhouette Score

Silhouette Score atau disebut juga sebagai koefisien *silhouette* merupakan suatu metrik yang digunakan untuk menilai seberapa baik hasil pengelompokan dalam analisis data. Metrik ini mengevaluasi sejauh mana sebuah data sesuai dengan klasternya sendiri dibandingkan dengan klaster lain (Atira & Sari, 2023).

2.9 Penelitian Terkait

Berikut adalah beberapa penelitian terkait mengenai topik penelitian yang dilakukan:

Penelitian terkait

Tabel 2. 1 Penelitian terkait

Judul	Nama Peneliti	Tahun	Hasil yang Didapatkan
Perbandingan Antara Algoritma <i>K-Means</i> dan Algoritma <i>Bisecting K-Means</i> Dalam Menganalisis Gempa Bumi di Indonesia	Adela Tania, Teny Handhayani, Janson Hendryli	2023	Hasil dari penelitian ini menunjukkan bahwa dalam menganalisis data gempa bumi, algoritma <i>K-Means</i> menunjukkan performa yang lebih unggul dibandingkan dengan algoritma <i>Bisecting K-Means</i> .

Komparasi Algoritma <i>K-Means</i> Dengan <i>K-Medoids</i> Dalam Klasterisasi Wilayah Rawan Bencana di Kabupaten Situbondo	Ganang Aji Pambudhi, Ahmad Homaidi, Firman Santoso	2024	Perbandingan hasil menunjukkan bahwa algoritma <i>K-Means</i> menghasilkan nilai DBI yang lebih baik dibandingkan dengan algoritma <i>K-Medoids</i> . Keunggulan ini terlihat dari nilai DBI <i>K-Means</i> yang lebih rendah, yaitu sebesar 0,853, meskipun menggunakan jumlah cluster dan dataset yang sama.
Pengelompokan Daerah Rawan Bencana Tanah Longsor di Provinsi Daerah Istimewa Yogyakarta Menggunakan Metode <i>K-Means Clustering</i>	Dwi Kartika Widyawati, Akhmad Fauzy	2024	Berdasarkan hasil analisis diperoleh 3 <i>cluster</i> yang terbentuk yaitu <i>cluster</i> 1 memiliki jumlah anggota sebanyak 13 Kecamatan. Kemudian <i>cluster</i> 2 dengan anggota sebanyak 5 Kecamatan. Kemudian <i>cluster</i> 3 dengan jumlah anggota sebanyak 60 Kecamatan.
Analisis Data Gempa Di Maluku Utara Menggunakan Algoritma <i>K-</i>	Isnawati Saleh, Gamaria Mandar, Junaidi Noh	2023	Grafik data gempa Maluku Utara tahun 2021 Visualisasi data gempa pada provinsi Maluku Utara tahun 2021 menunjukan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

<i>Means dan Liner Regression</i>			terdapat 852 dengan 11 titik sebaran lokasi kejadian gempa bumi dan Grafik data gempa Maluku Utara tahun 2022 Visualisasi data gempa pada provinsi Maluku Utara tahun 2022 menunjukkan terdapat 498 dengan 11 titik sebaran lokasi kejadian gempa bumi.
Penerapan Metode <i>Clustering</i> Untuk Pemetaan Daerah Rawan Bencana di Kabupaten Bojonegoro	Alif Yuanita Kartini, Deby Fakhriyana	2024	Pengelompokan wilayah rawan bencana di Kabupaten Bojonegoro dilakukan ke dalam empat cluster menggunakan metode K-Means sebagai metode yang paling optimal. Berdasarkan hasil analisis, dapat disimpulkan bahwa Kecamatan Bojonegoro merupakan daerah dengan tingkat kerawanan bencana tertinggi di kabupaten tersebut.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 3

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Metodologi penelitian berisi panduan proses tahapan penelitian yang dilaksanakan sehingga penelitian sesuai dengan tujuan penelitian. Tahapan penelitian pada penelitian ini adalah sebagai berikut.



3.2 Identifikasi Masalah

Tahapan awal yang dilakukan dalam penelitian ini adalah identifikasi masalah. Identifikasi masalah diperlukan untuk mengetahui masalah yang akan



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

diteliti. Penelitian yang dilakukan mengangkat masalah mengenai data gempa yang terjadi di Indonesia.

3.3 Studi Pustaka

Studi pustaka dilakukan untuk mencari informasi dari berbagai hal yang berhubungan dengan penelitian yang akan dilaksanakan. Informasi akan dicari melalui berbagai literatur seperti, jurnal, sumber-sumber dari internet dan sumber lainnya untuk mendapatkan pengetahuan mengenai hal yang akan diteliti. Pada penelitian ini, digunakan berbagai sumber informasi dari jurnal, *e-book*, dan situs internet yang berkaitan dengan topik penelitian yang dilakukan.

3.4 Perumusan Masalah

Pada langkah ini, dilakukan perumusan masalah terhadap penelitian yang dilakukan. Perumusan masalah dilakukan setelah melakukan identifikasi terhadap masalah dan dilakukannya studi literatur terhadap penelitian yang akan dilaksanakan. Perumusan masalah dalam penelitian ini adalah bagaimana menerapkan algoritma *K-Means* untuk klasterisasi data gempa yang terjadi di Indonesia.

3.5 Pengumpulan Data

Setelah perumusan masalah selesai ditentukan, maka dimulailah proses pengumpulan data yang berkaitan dengan penelitian. Data yang digunakan dalam penelitian ini adalah data sekunder bersumber dari situs kaggle.

3.6 Analisis Data

Pada tahap analisis data dilakukan proses pengolahan data untuk mendapatkan identifikasi pola, hubungan, serta informasi penting yang terdapat didalamnya. Hal ini dilakukan untuk bisa mendapatkan pemahaman yang lebih mendalam mengenai data yang dianalisis dan mengambil keputusan berdasarkan informasi yang ditemukan. Pada penelitian ini, akan dilakukan tahapan-tahapan pada *Knowledge Discovery in Database (KDD)*. Data akan diolah dengan menggunakan algoritma *K-Means Clustering*, dengan menggunakan penghitungan *Euclidean Distance* dan *Manhattan Distance*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.6.1 Analisis Proses KDD

Pada tahapan ini dilakukan proses analisis tahapan-tahapan proses KDD untuk klusterisasi data gempa di Indonesia menggunakan *K-Means Clustering*. Tahapan KDD yang dilakukan adalah sebagai berikut.

1. Data Selection

Tahapan ini dilakukan untuk mengambil data dari dataset yang diperlukan untuk diolah menggunakan *K-Means Clustering*.

2. Data Preprocessing

Tahapan ini dilakukan untuk menghindari data dari data-data yang mengganggu seperti data kosong, duplikat, dan data yang tidak konsisten.

3. Transformation

Pada tahapan ini, dilakukan transformasi data menjadi data yang sesuai dengan model ataupun algoritma *K-Means* yang digunakan dalam mengolah data.

4. Data Mining

Pada tahapan ini, dilakukan proses pencarian dan penggalian pengetahuan sehingga diperoleh sebuah model yang bisa dijadikan sebagai sebuah informasi penting dan berguna menggunakan *K-means Clustering*. Penghitungan *K-Means Clustering* yang dilakukan menggunakan 2 penghitungan jarak, yaitu *Euclidean Distance* dan *Manhattan Distance*.

- a. Euclidean Distance

Dilakukan penghitungan dengan menggunakan *Euclidean Distance*. Hasil *cluster* yang didapatkan akan divisualisasikan untuk melihat persebaran data pada tiap *cluster* yang dihasilkan. Didapatkan jumlah data pada tiap *cluster* yang terbentuk beserta wilayah-wilayah yang termasuk ke dalam tiap *cluster*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. Manhattan Distance

Kemudian dilakukan juga penghitungan *K-Means Clustering* dengan menggunakan *Manhattan Distance*. Hasil penghitungan akan dilakukan visualisasi data untuk melihat persebaran data *cluster*. Didapatkan jumlah data pada *cluster* yang terbentuk dan juga wilayah-wilayah yang masuk ke *cluster* tersebut.

5. Evaluation

Ini merupakan tahapan akhir pada proses KDD. Dilakukan evaluasi hasil dari *K-Means Clustering* yang dilakukan. Dilakukan penghitungan *Elbow Method* untuk menilai jumlah *cluster* optimum pada pengolahan data. Kemudian dilakukan penghitungan *Silhouette Score* pada penghitungan *Euclidean Distance* dan *Manhattan Distance* untuk menilai penghitungan jarak mana yang lebih baik dalam klasterisasi data gempa di Indonesia menggunakan *K-Means Clustering*. Nilai *Silhouette Score* dinilai lebih baik jika semakin mendekati 1.

3.7 Kesimpulan dan Saran

Pada tahapan ini, hasil dari penelitian akan dirangkum dan disimpulkan. Pada tahapan saran, akan berisi saran dan kritik membangun untuk membantu pada penelitian berikutnya dan melakukan perbaikan atas kesalahan dan kekurangan dalam penelitian. Selain itu, juga diberi ajakan untuk pengembangan lebih lanjut dari penelitian yang dilakukan dan eksplorasi lebih lanjut berhubungan dengan penelitian yang dilakukan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB 5

PENUTUP

5.1 Kesimpulan

Dari penelitian yang dilakukan, didapatkan bahwa algoritma *K-Means* dapat diterapkan dalam klasterisasi data gempa di Indonesia.

Pada penghitungan *K-Means Clustering* dengan menggunakan *Euclidean Distance* didapatkan *cluster* 1 memiliki 49838 *record*, *cluster* 2 berisi 33126 *record*, dan *cluster* 3 dengan 7188 *record*. *Cluster* 1 memiliki karakteristik kedalaman dan magnitudo gempa yang rendah pada data penelitian. Wilayah yang paling banyak muncul pada *cluster* 1 adalah *Minahassa Peninsula* (Semenanjung Minahasa), Sulawesi, Sumbawa, Jawa (Jawa), dan *Northern Sumatra* (Utara Sumatera). *Cluster* 2 memiliki kedalaman gempa yang rendah dan magnitudo yang cukup tinggi. Wilayah yang paling banyak muncul pada *cluster* 2 adalah *Northern Molucca Sea* (Laut Maluku Utara), Semenanjung Minahasa, *Southern Sumatra* (Selatan Sumatera), Irian Jaya, dan *Talaud Islands* (Pulau Talaud). *Cluster* 3 memiliki karakteristik kedalaman gempa yang tinggi dengan magnitudo sedang. Wilayah yang paling banyak muncul pada *cluster* 3 adalah *Banda Sea* (Laut Banda), Semenanjung Minahasa, Flores, Halmahera, dan *Celebes Sea* (Laut Celebes).

Pada penghitungan *K-Means Clustering* dengan menggunakan *Manhattan Distance* didapatkan *cluster* 1 memiliki 36525 *record*, *cluster* 2 dengan 46676 *record*, dan *cluster* 3 berisikan 6951 *record*. *Cluster* 1 merupakan daerah dengan karakteristik kedalaman yang rendah dan magnitudo tinggi. Wilayah yang paling banyak muncul pada *cluster* 1 adalah *Northern Molucca Sea* (Laut Maluku Utara), Semenanjung Minahasa, *Southern Sumatra*, Jawa, dan Irian Jaya. *Cluster* 2 memiliki karakteristik kedalaman rendah dengan magnitudo rendah. Wilayah yang paling banyak muncul pada *cluster* 2 adalah *Minahassa Peninsula* (Semenanjung Minahasa), Sulawesi, Sumbawa, Jawa, dan *Northern Sumatra*. *Cluster* 3 memiliki karakteristik dengan kedalaman tinggi dan magnitudo sedang. Wilayah yang paling

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

banyak muncul pada *cluster* 3 adalah *Banda Sea* (Laut Banda), Semenanjung Minahasa, Flores, Halmahera, dan Laut Celebes.

Setelah dilakukan penghitungan dengan *Elbow Method*, didapatkan nilai optimal adalah 3. Didapatkan nilai *Silhouette Score* pada penghitungan *Euclidean Distance* penelitian ini adalah 0,469702780513585. Pada penghitungan dengan *Manhattan Distance*, didapatkan nilai *Silhouette Score* yaitu 0,44916422524456295. Dapat disimpulkan bahwa nilai *Silhouette Score* pada penghitungan *Euclidean Distance* menghasilkan kualitas *clustering* yang lebih baik dari penghitungan *Manhattan Distance* pada penelitian klasterisasi data gempa di Indonesia menggunakan *K-Means Clustering* ini.

5.2 Saran

Hasil *Silhouette Score* yang didapat pada penghitungan *Euclidean Distance* dan *Manhattan Distance* terhitung cukup baik. Diharapkan data dapat diolah dengan algoritma *clustering* lainnya ataupun penghitungan dengan jarak lainnya untuk dapat mencari kualitas *clustering* yang lebih baik untuk penghitungan klasterisasi data gempa di Indonesia.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Aghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 9(2), 75–81.
- Atira, A., & Sari, B. N. (2023). Penerapan Silhouette Coefficient, Elbow Method dan Gap Statistics untuk Penentuan Cluster Optimum dalam Pengelompokan Provinsi di Indonesia Berdasarkan Indeks Kebahagiaan. *Jurnal Ilmiah Wahana Pendidikan*, 9(17), 76–86. <https://doi.org/https://doi.org/10.5281/zenodo.8282638>
- Dosbing. (2024). Implementasi dan Perhitungan Manual Algoritma K-Means dalam Clustering Data. Diakses pada 8 Mei 2025, dari <https://dosbing.id/2024/06/24/implementasi-dan-perhitungan-manual-algoritma-k-means-dalam-clustering-data/>
- Duha, T., Laia, M., Huda, A. K., & Jasuma, A. (2023). Klasifikasi Data Gempa Bumi di Pulau Sumatera Menggunakan Algoritma Naïve Bayes. *Jurnal Informatika*, 2(1), 23–27. <https://doi.org/10.57094/ji.v2i1.840>
- Fitri, T. I., Sari, H. L., Dinata, R. K., Hasdyna, N., Hasugian, B. S., Retno, S., ... Pohan, N. (2024). *Data Mining*. Payakumbuh: Serasi Media Teknologi.
- Heraldi, H. Y., Aprilia, N. C., & Pratiwi, H. (2019). Analisis Cluster Intensitas Kebencanaan di Indonesia Menggunakan Metode K-Means. *Indonesian Journal of Applied Statistics*, 2(2), 137. <https://doi.org/10.13057/ijas.v2i2.34911>
- Musid, F. N., Dikananda, A. R., & Fathurrohman. (2023). Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Data Jumlah Kerusakan Rumah Berdasarkan Kondisi di Jawa Barat. *Journal of Student Research (JSR)*, 1(3), 101–114.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Nurhayati. (2022). *Pemodelan K-Means Algoritma dan Big Data Analysis (Pemetaan Data Mustahiq)*. Jakarta: Pascal Books.

Oktavian, C. N., & Adiarto, A. (2017). *Mitigasi Bencana*. Ponorogo: Wade Group.

Pambudhi, G. A., Homaidi, A., & Santoso, F. (2024). Komparasi Algoritma K-Means Dengan K-Medoids Dalam Klasterisasi Wilayah Rawan Bencana Di Kabupaten Situbondo. *Jurnal Teknik Elektro Dan Informatika*, 19(2), 173–179.

Qothrunnada, N. H., Utami, R. Y., & Rizky, S. A. (2022). Menganalisis bencana alam gempa bumi dalam perspektif Al-Quran. *Jurnal Konferensi Integrasi Interkoneksi Islam Dan Sains*, 4, 257–260.

Ruyani. (2023). *Seri Fenomena Alam dan Mitigasi : Gempa Bumi*. Jakarta: Bumi Aksara.

Safira, R., & Castaka, A. S. (2024). Optimalisasi Algoritma K-Means Menggunakan Metode Elbow Dalam Pengelompokan Data Stunting. *Brahmana : Jurnal Penerapan Kecerdasan Buatan*, 5(2), 257–264. Retrieved from <https://tunasbangsa.ac.id/pkm/index.php/brahmana/article/view/396>

Saleh, I., Mandar, G., & Noh, J. (2023). Analisis Data Gempa Di Maluku Utara Menggunakan Algoritma K-Means Dan Linear Regression. *Dintek*, 16(2), 12–19.

Setiawan, I. N., Krismawati, D., Pramana, S., & Tanur, E. (2022). Klasterisasi Wilayah Rentan Bencana Alam Berupa Gerakan Tanah Dan Gempa Bumi Di Indonesia. *Seminar Nasional Official Statistics*, 2022(1), 669–676. <https://doi.org/10.34123/semnasoffstat.v2022i1.1538>

Setiawan, Z., Fajar, M., Priyatno, A. M., Putri, A. Y. P., Aryuni, M., Yuliyanti, S., ... Wijaya, A. (2023). *Buku Ajar Data Mining*. Jambi: PT. Sonpedia Publishing Indonesia.

Sutirta, N. F., & Noviandi. (2024). Perbandingan Manhattan dan Euclidean

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Distance Untuk Pengelompokan Penyakit Jantung Menggunakan Algoritma K-Means. *ICIT Journal*, 10(1), 61–70.

<https://doi.org/10.33050/icit.v10i1.2860>

Lania, A., Handhayani, T., & Hendryli, J. (2023). Perbandingan Antara Algoritma K-Means Dan Algoritma Bisecting K-Means Dalam Menganalisis Gempa Bumi Di Indonesia. *Simtek : Jurnal Sistem Informasi Dan Teknik Komputer*, 8(2), 265–270. <https://doi.org/10.51876/simtek.v8i2.205>

Widyawati, D. K., & Fauzy, A. (2024). Pengelompokan Daerah Rawan Bencana Tanah Longsor di Provinsi Daerah Istimewa Yogyakarta Menggunakan Metode K-Means Clustering. *Emerging Statistics and Data Science Journal*, 2(2), 212–221.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR RIWAYAT HIDUP

Nama	: Hayyudi Giannanda Asfa
NIM	: 11850112515
Program Studi	: Teknik Informatika
Email	: 11850112515@students.uin-suska.ac.id
Riwayat Pendidikan	
2006-2012	: SD Negeri 006 Rengat
2012-2015	: SMP Negeri 1 Rengat
2015-2018	: SMA Negeri 1 Rengat
2018-2025	: S1 Teknik Informatika UIN Sultan Syarif Kasim Riau

UIN SUSKA RIAU