



## IMPLEMENTASI RAPIDMINER PADA KLASTERISASI GEMPA BUMI DI INDONESIA BERDASARKAN KEDALAMAN MENGGUNAKAN K-MEANS

Ma'ruf Ubaidillah<sup>a\*</sup>, Zaehol Fatah<sup>b</sup>

<sup>a</sup>Prodi Teknologi Informasi, Fakultas Sains dan Teknologi, [mrfubaidillah@gmail.com](mailto:mrfubaidillah@gmail.com), Universitas Ibrahimy Sukorejo  
Situbondo Jawa Timur

<sup>b</sup>Prodi Sistem Informasi, Fakultas Sains dan Teknologi, [zaeholfatah@gmail.com](mailto:zaeholfatah@gmail.com), Universitas Ibrahimy Sukorejo  
Situbondo Jawa Timur

\* Korespondensi

### ABSTRACT

*Indonesia is a meeting place for four major tectonic plates, namely the Carolina Plate, the Philippine Sea Plate, the Indo-Australian Plate, and the Eurasian Plate. Every year thousands of earthquakes strike, causing material losses, infrastructure damage, and even loss of life. Therefore, understanding the characteristics of earthquakes in Indonesia is a crucial step to mitigate disaster risks and improve community preparedness. The dataset used comes from the Kaggle.com website, the dataset is taken from the Earthquake Repository managed by BMKG. The K-Means algorithm is used as a clustering process method in this study. Clustering using K-Means aims to identify the dominant types of earthquakes that occur in regions of Indonesia. The application of this method to earthquakes that occurred in Indonesia based on the identified depth in cluster\_0 shows the least earthquakes with the deepest earthquake type. Cluster\_1 shows a shallow earthquake type. While cluster\_2 is the earthquake with the most occurrences and shows an earthquake with a moderate depth.*

**Keywords:** Earthquake, Rapidminer, Clustering, Data Mining, K-Means.

### Abstrak

Indonesia menjadi tempat bertemunya empat lempeng tektonik utama yaitu Lempeng Carolina, Lempeng Laut Filipina, Lempeng Indo-Australia, dan Lempeng Eurasia. Setiap tahunnya ribuan gempa bumi melanda, membawa dampak berupa kerugian material, kerusakan infrastruktur, hingga jatuhnya korban jiwa. Oleh sebab itu memahami karakteristik gempa bumi di Indonesia menjadi langkah krusial untuk memitigasi risiko bencana dan meningkatkan kesiapsiagaan masyarakat. Dataset yang digunakan berasal dari situs web Kaggle.com, dataset tersebut diambil dari Repositori Gempa Bumi yang dikelola oleh BMKG. Algoritma K-Means digunakan sebagai metode proses klusterisasi pada penelitian ini. Klusterisasi dengan menggunakan K-Means bertujuan untuk mengidentifikasi tipe gempa yang dominan terjadi di wilayah-wilayah Indonesia. Penerapan metode ini pada gempa bumi yang terjadi di Indonesia berdasarkan kedalaman teridentifikasi pada cluster\_0 menunjukkan gempa paling sedikit terjadi dengan tipe gempa paling dalam. Pada cluster\_1 menunjukkan tipe gempa dangkal. Sedangkan pada cluster\_2 menjadi gempa dengan jumlah paling banyak terjadi dan menunjukkan gempa dengan kedalaman sedang.

**Kata Kunci:** Gempa Bumi, Rapidminer, Klusterisasi, Data Mining, K-Means

### 1. PENDAHULUAN

Saat ini, secara umum Indonesia berada pada interaksi empat lempeng utama, yakni Lempeng Eurasia yang memuat Indonesia Barat (Paparan Sunda), Lempeng Indo-Australia yang mencakup Indonesia Timur, serta Lempeng Laut Filipina dan Lempeng Karolina di sisi utara Indonesia Timur. Keempat lempeng tersebut berinteraksi dengan cara yang berbeda-beda[1]. Setiap tahunnya ribuan gempa bumi melanda, membawa dampak berupa kerugian material, kerusakan infrastruktur, hingga jatuhnya korban jiwa. Oleh sebab itu

memahami karakteristik gempa bumi di Indonesia menjadi langkah krusial untuk memitigasi risiko bencana dan meningkatkan kesiapsiagaan masyarakat.

Gempa bumi dengan kedalaman episenter (hiposentrum) kurang dari 50 km disebut gempa dangkal. Kerusakan yang besar biasa ditimbulkan ketika gempa dangkal ini disertai dengan kekuatan besar. Gempa bumi terbanyak 85% merupakan gempa dangkal. Gempa bumi yang mempunyai kedalaman hiposenter antara 50-300 km disebut gempa sedang. Gempa jenis ini hanya terjadi pada 12% dari seluruh gempa tektonik. Gempa bumi yang mempunyai kedalaman hiposenter lebih dari 300 km disebut gempa dalam. Gempa jenis ini hanya mencakup 3% dari seluruh gempa tektonik yang termasuk dalam kategori gempa dalam[2].

Dengan berkembangnya zaman, kemajuan teknologi dalam proses pengumpulan dan penyimpanan data yang cepat serta akurat memungkinkan organisasi untuk mengumpulkan data dalam jumlah besar. Oleh karena itu, penggunaan metode manual untuk menganalisis data tidak lagi efektif dalam mengekstraksi informasi dari volume data yang sangat besar. Teknologi baru yang dapat membantu organisasi atau perusahaan adalah data mining, teknologi tersebut berguna untuk menggali informasi yang dibutuhkan dari gudang data yang mereka miliki[3]. Data mining merupakan sebuah teknologi yang dapat mengolah data dalam jumlah besar yang digunakan oleh perusahaan untuk mengubah data mentah menjadi informasi yang berguna untuk mengambil keputusan bisnis yang sangat penting[4]. Data mining adalah proses untuk menemukan informasi dan pola yang bernilai dari kumpulan data dalam jumlah besar. Proses ini mencakup pengumpulan, ekstraksi, analisis, hingga pengolahan data secara statistik. Data mining juga sering disebut sebagai Knowledge Discovery, Knowledge Extraction, Data/Pattern Analysis, Information Harvesting, dan istilah lainnya.[5].

Analisis klusterisasi berdasarkan kedalaman gempa dapat membantu dalam mengidentifikasi tipe gempa dominan yang terjadi di berbagai wilayah Indonesia. Salah satu metode yang efektif untuk melakukan klusterisasi adalah K-Means, K-Means Clustering adalah teknik pengelompokan data non-hirarki yang memisahkan data ke dalam cluster, mengelompokkan data dengan fitur yang sama bersama-sama dan mengelompokkan data dengan karakteristik yang berbeda ke dalam kelompok yang berbeda[6]. Dalam algoritma K-means, setiap data harus termasuk ke cluster tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke cluster yang lain. Pada dasarnya penggunaan algoritma K-means dalam melakukan proses clustering tergantung dari data yang ada dan konklusi yang ingin dicapai[7]. K-means clustering bekerja dengan membagi data menjadi beberapa cluster tergantung pada kesamaan pola fitur yang digunakan[8]. Sehingga dengan metode ini dapat mengelompokkan data gempa berdasarkan kedalaman dan pola distribusi di setiap wilayah terlihat lebih jelas. Dan dengan metode ini, dapat mengelompokkan gempa menjadi beberapa cluster tipe-tipe gempa yang dominan terjadi, sehingga membantu dalam pemahaman karakteristik dan penyebaran tipe-tipe gempa di Indonesia berdasarkan kedalamannya.

Penelitian sebelumnya di wilayah Sulawesi sering mengalami kerusakan akibat gempa bumi berkekuatan besar yang terjadi pada kedalaman tertentu. Salah satu faktor yang memengaruhi kerusakan, selain kekuatan gempa, adalah hiposentrum gempa. Algoritma Cluster machine learning K-Means digunakan untuk memeriksa kedua faktor tersebut guna menemukan pola kekuatan gempa menurut kedalaman. Visualisasi cluster menunjukkan bahwa gempa bumi yang lebih besar dari atau sama dengan 6 SR sering terjadi pada kedalaman dangkal. Hal ini menunjukkan bagaimana metode cluster machine learning K-Means dapat mengidentifikasi pola kedalaman kejadian gempa bumi dan bagaimana beberapa provinsi di Sulawesi dalam cluster pertama dapat terdampak oleh bahaya gempa bumi[9].

Pada penelitian lain, penggunaan informasi tentang magnitudo, kedalaman, dan waktu gempa bumi untuk menganalisis pola aktivitas seismik di Pulau Sumatera. Metodologi analisis yang digunakan dalam penelitian ini terdiri dari pengelompokan data menggunakan teknik K-Means Clustering dan mendeteksi outlier menggunakan visualisasi boxplot. Selain itu, model Rantai Markov dibuat menggunakan temuan pengelompokan untuk memprediksi pola gempa bumi berdasarkan perubahan keadaan. Kedalaman sedang hingga sedang (0,7165), durasi awal tahun hingga awal tahun (0,9942), dan magnitudo sedang hingga sedang (0,5339) merupakan peluang transisi terbesar, menurut temuan utama. Di Pulau Sumatera, metode ini secara signifikan mendukung perencanaan tanggap gempa bumi dan langkah-langkah pengurangan risiko bencana[10]. menunjukkan metode pengelompokan seperti pengelompokan K-Means efektif untuk analisis dan visualisasi pada data seismik yang rumit.

Klasterisasi perlu dilakukan upaya untuk pengelompokan wilayah di Indonesia berdasarkan gempa bumi yang sering terjadi agar dapat memudahkan proses mitigasi bencana[11]. Oleh karena itu proses clustering gempa di Indonesia dengan menggunakan metode K-Means adalah upaya untuk menentukan gempa dominan yang terjadi berdasarkan kedalaman menjadi tujuan dalam penelitian ini. Hasil dari penelitian ini diharapkan dapat memberikan informasi yang bermanfaat bagi peneliti-peneliti selanjutnya, terutama dalam usaha mitigasi dan perencanaan tanggap bencana yang lebih efektif

## 2. TINJAUAN PUSTAKA

### 2.1. Algoritma K-Means

Algoritma K-means merupakan metode klasterisasi dengan membagi atau memisahkan data ke dalam beberapa kategori untuk mengelompokkan data. K-means dapat mengurangi jarak rata-rata antara setiap titik data dan klusternya melalui pemisahan berulang. Setiap bagian data dalam algoritma K-means harus menjadi bagian dari kluster tertentu pada titik waktu tertentu; setelah itu, data tersebut dapat berpindah ke kluster lain. Intinya, penggunaan algoritma K-means dalam proses pengelompokan ditentukan oleh data yang tersedia saat ini dan hasil yang diinginkan.[7]. Untuk itu digunakan algoritma K-means yang didalamnya memuat aturan sebagai berikut:

- a. Menginputkan jumlah cluster
- b. Hanya memiliki atribut bertipe numerik

Subset dari banyak komponen populasi pertama kali dipilih menggunakan algoritma K-means untuk berfungsi sebagai pusat kluster pertama. Pusat kluster dipilih secara acak dari kumpulan populasi data yang sudah ada sebelumnya pada tahap ini. Setiap komponen dalam populasi data kemudian dikenakan pengujian K-means, yang menugaskannya ke salah satu pusat kluster yang ditunjuk berdasarkan jarak terkecil antara komponen tersebut dan setiap pusat kluster. Setelah semua komponen data dikategorikan ke dalam setiap kluster, lokasi pusat kluster akan dihitung ulang, dan akhirnya posisi kluster baru akan dibuat.[7].

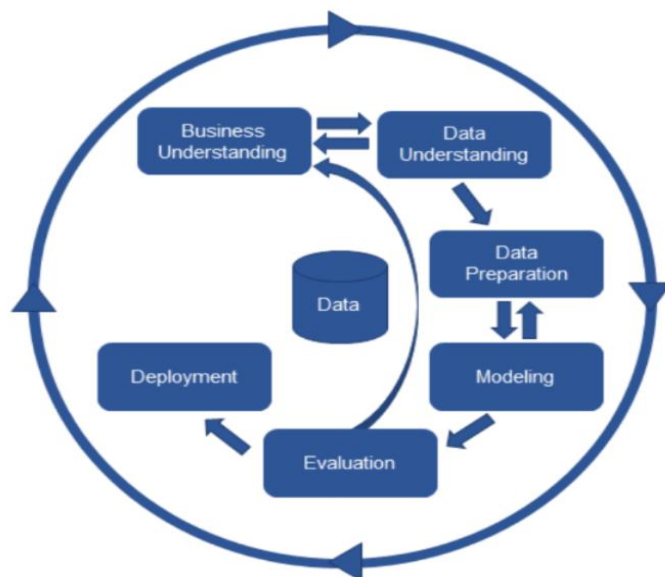
### 2.2. Rapid Miner

Rapid Miner adalah software yang dibuat oleh Dr. Markus Hofmann dari Institut Teknologi Blanchardstown dan Rafi Klinkenberg dari rapid-i.com, yang memiliki antarmuka pengguna grafis (GUI) yang intuitif[12]. Program RapidMiner perangkat lunak yang bersifat sumber terbuka. RapidMiner adalah alat untuk analisis prediktif, penambangan teks, dan penambangan data. Untuk membantu pengguna membuat keputusan terbaik, RapidMiner menggunakan berbagai teknik deskriptif dan prediktif. Ada lebih dari 500 operator penambangan data di RapidMiner, termasuk operator untuk input, output, persiapan data, dan visualisasi. RapidMiner adalah mesin penambangan data yang dapat dihubungkan ke produk lain dan merupakan program mandiri untuk analisis data. Karena RapidMiner ditulis dalam Java, program ini kompatibel dengan semua sistem operasi.[13].

Fokus lebih lanjut dari Rapid Miner konektivitas tinggi ke berbagai sumber data terluas, termasuk Oracle, IBMDB2, Microsoft SQL Server, MySQL, Postgre SQL, dan Ingres, serta akses ke Excel, Access, dan SPSS, di antara format data lainnya, merupakan area fokus lain untuk Rapid Miner. Dengan demikian, Rapid Miner dapat digunakan untuk integrasi dan transformasi data (ETL) selain analisis data dengan hasil yang luar biasa, bersama dengan ratusan operator untuk pra-pemrosesan data. Selain itu, pelanggan dapat memilih dari berbagai format perangkat lunak.[12].

## 3. METODOLOGI PENELITIAN

CRISP-DM digunakan sebagai metodologi pada penelitian ini. NCR Systems Engineering Copenhagen, DaimlerChrysler AG, SPSS Inc., dan OHRA Verzekeringen en Bank Groep B.V. membentuk kolaborasi untuk menciptakan teknik penambangan data CRISP-DM. Enam langkah membentuk model proses analisis data yang ditetapkan dalam CRISP-DM. Dalam CRISP-DM, tahapan diselesaikan secara berurutan, dengan opsi untuk kembali ke langkah sebelumnya jika masalah muncul selama tahap saat ini.[14]. Tahapan CRISP-DM dapat dilihat pada gambar dibawah ini.



Gambar 1. Tahapan CRISP-DM

Sumber : Dody Dharma Hutabarat dkk, 2021 [14]

### 3.1 Business Understanding

Tahap ini menguraikan definisi masalah, tujuan penelitian, dan pembicaraan terkait proses bisnis. Pada tahap ini, penting untuk mengonfirmasi tujuan unit kerja atau masalah yang dihadapi. Sumber daya, data, risiko, pengeluaran, dan imbalan semuanya harus direncanakan dalam inisiatif analisis data. Seiring dengan perencanaan dan tujuan proyek, manajemen proyek juga diputuskan[14].

### 3.2 Data understanding

Memahami data dan memilih data yang mendukung tujuan penelitian merupakan tujuan utama dari tahap ini. Langkah ini melibatkan pengumpulan data awal, analisis data yang *diperoleh*, pemeriksaan data, konfirmasi kualitas data, pemahaman keadaan data yang menyimpang atau outlier, prediksi korelasi antar variabel, dan penyajian data[14].

### 3.3 Data preparation

Pemilihan atribut yang dapat diatur untuk membuat kumpulan data untuk pemodelan merupakan bagian dari langkah ini. Langkah ini mencakup aktivitas berikut: memilih data mana yang akan *digunakan*, membersihkan data yang tidak akan digunakan, membuat variabel baru jika perlu, dan mengintegrasikan data[14].

### 3.4 Modeling

Pada tahap ini diterapkan teknik atau metodologi data mining guna memperoleh pemodelan yang sesuai dengan tujuan penelitian. Pada tahap ini, sejumlah tugas harus diselesaikan, antara lain memilih model yang akan digunakan, mengumpulkan atau mengatur data untuk dijadikan sampel, membuat model, mengevaluasi model untuk mengidentifikasi faktor-faktor signifikan, dan berbicara dengan unit pemilik proses bisnis[14].

### 3.5 Evaluation

Tujuan dari langkah ini adalah untuk menilai pemodelan yang telah dilakukan guna memenuhi tujuan studi. Di antara tugas-tugas yang harus diselesaikan adalah menilai hasil keseluruhan untuk melihat apakah *tujuan* proyek telah tercapai, menilai aspek teknis dan praktis dari model, seperti kecepatan dan kinerja, memeriksa prosedur yang diikuti, dan memutuskan apakah model akhir akan dilanjutkan ke tahap penyebaran atau tidak[14].

### 3.6 Deployment

Pada langkah ini menunjukkan bahwa pemodelan siap digunakan dan sesuai dengan tujuan studi. Pengambilan keputusan tentang rencana implementasi, pemantauan dan *pemeliharaan*, serta pembuatan

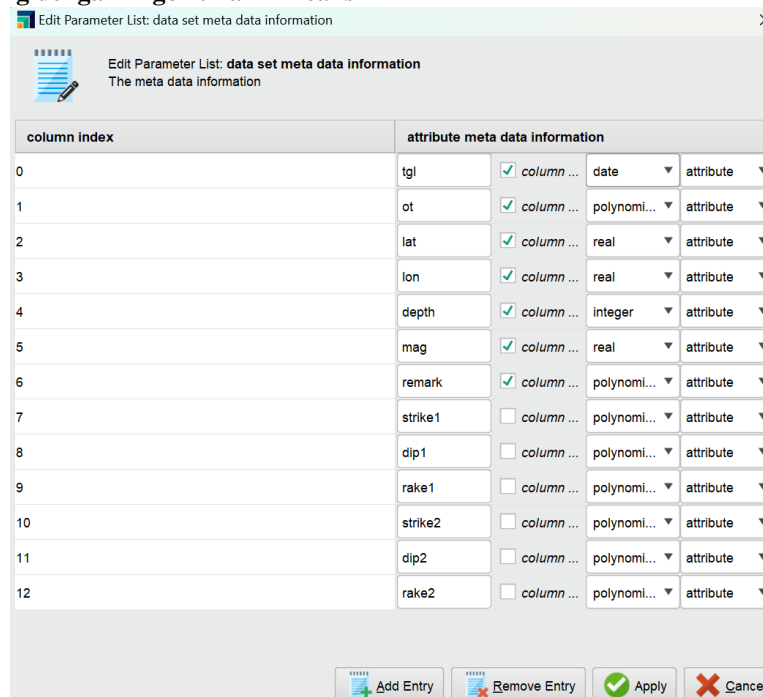
dokumen proyek akhir merupakan beberapa tindakan yang harus diambil jika temuan penilaian menunjukkan bahwa model siap untuk dunia produksi.[14].

#### 4. HASIL DAN PEMBAHASAN

##### 4.1 Pemahaman Data

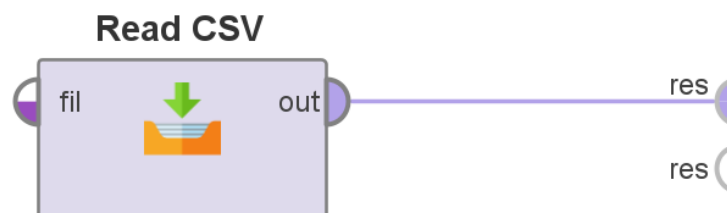
Data yang akan di proses berasal dari website Kaggle.com, dataset tersebut di ambil dari Repositori Gempa Bumi yang dikelola oleh BMKG (Badan Meteorologi, Klimatologi, dan Geofisika) yang merupakan lembaga pemerintah non-departemen Indonesia, dataset dapat di unduh pada url berikut : <https://www.kaggle.com/datasets/kekavigi/earthquakes-in-indonesia>. Pada dataset tersebut memiliki 9 kolom terdiri dari : tgl (tanggal kejadian), ot (cap waktu kejadian), lat (lintang episentrum kejadian (derajat), berkisar antara 6N hingga 11S), lon ( garis bujur episentrum kejadian (derajat), berkisar antara 142E hingga 94E), depth (kedalaman kejadian (km) mag: Besarnya peristiwa berkisar antara 1 sampai 9,5), remark (Wilayah Flinn-Engdahl dari kejadian tersebut), dan beberapa kolom lainnya seperti nilai dip1, strike1, rake1, dip2, strike2, dan rake2.

##### 4.2 Proses Clustering dengan Algoritma K-Means



Gambar 2. Seleksi dataset

Pada tahap ini data yang didapat dari Kaggle di import kedalam rapid miner. Pastikan file berbentuk .CSV, setelah di import data tersebut diseleksi hanya mengambil data-data yang diperlukan saja seperti : tgl,ot,lat,lon,depth, mag, remark.



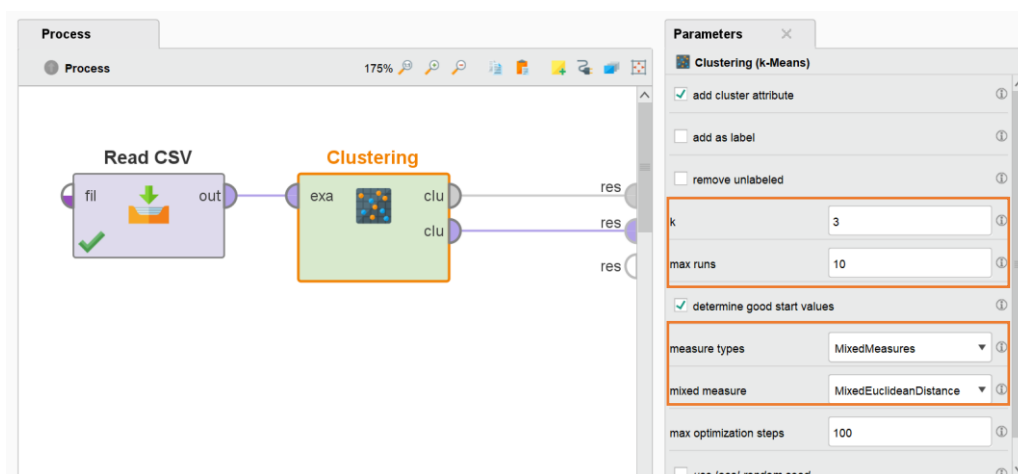
Gambar 3. Baca File CSV

Setelah proses seleksi data selesai, kemudian dilakukan uji coba dataset yang nantinya akan di klusterisasi menggunakan algoritma K-Means.

Row No.	tgl	ot	lat	lon	depth	mag	remark
1	Nov 1, 2008	21:02:43.058	-9.180	119.060	10	4.900	Sumba Regio...
2	Nov 1, 2008	20:58:50.248	-6.550	129.640	10	4.600	Banda Sea
3	Nov 1, 2008	17:43:12.941	-7.010	106.630	121	3.700	Java - Indone...
4	Nov 1, 2008	16:24:14.755	-3.300	127.850	10	3.200	Seram - Indo...
5	Nov 1, 2008	16:20:37.327	-6.410	129.540	70	4.300	Banda Sea
6	Nov 1, 2008	14:47:00.029	-7.370	105.310	18	3.300	Java - Indone...
7	Nov 1, 2008	13:04:38.742	0.100	98.550	12	4.700	Northern Su...
8	Nov 1, 2008	10:23:51.646	-7.070	129.670	135	4.800	Banda Sea
9	Nov 1, 2008	09:50:32.503	-3.320	128.020	10	2.300	Seram - Indo...
10	Nov 1, 2008	06:50:52.220	-4.430	127.450	10	3.200	Banda Sea
11	Nov 1, 2008	06:42:18.915	-3.940	127.450	9	2.600	Seram - Indo...
12	Nov 1, 2008	06:23:00.287	-3.430	128.390	10	2.300	Seram - Indo...
13	Nov 1, 2008	06:01:05.398	-4.260	127.660	10	3.200	Banda Sea

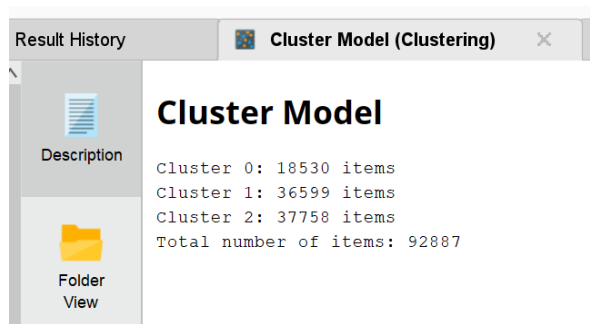
Gambar 4. Dataset gempa bumi

Hasil data yang telah di jalankan menunjukkan bahwa ada 92,887 data record yang bisa di olah. Kemudian clustering dataset tersebut menggunakan algoritma K-Means, dengan menambahkan operator clustering K-Means pada Rapid Miner.



Gambar 5. Proses Clustering

Untuk melakukan clustering menggunakan k-means, data di proses dengan parameter k : 3 yang berarti cluster yang akan digunakan sebanyak 3 cluster saja, max runs : 10 yang berarti K-Means melakukan proses clustering sebanyak sepuluh kali, dan measure types : MixedMeasures karena dataset yang digunakan memiliki atribut yang berbeda-beda.



Gambar 6. Cluster model

Hasil dari clustering menggunakan algoritma K-Means menunjukkan cluster\_0 : 18,530 items, cluster\_1 : 36,599 items, dan cluster\_2 : 37,758 items.

Attribute	cluster_0	cluster_1	cluster_2
tgl	1313164133621.155	1621316935555.616	1506382071571.587
ot	9247.866	72481.933	36678.319
lat	-3.133	-3.337	-3.603
lon	118.077	119.238	119.615
depth	57.256	45.520	48.345
mag	3.890	3.459	3.577
remark	13.604	14.360	14.560

Gambar 7. Centroid table

Untuk mengetahui mana cluster yang dangkal, sedang, dan dalam dari setiap cluster dengan melihat nilai centroid pada atribut depth pada masing-masing cluster bisa dilihat pada Centroid Table (gambar 7). Cluster\_0 menunjukkan nilai tertinggi 57,256 yang berarti gempa dalam, sementara cluster\_1 dengan nilai centroid terendah 45,520 yang berarti gempa dangkal, dan cluster\_2 dengan nilai 48,345 yang berarti gempa dengan kedalaman sedang.

Proses diatas menghasilkan data yang informatif, tapi data tersebut belum ringkas. Oleh karena itu di rangkum agar dapat menghasilkan data yang mudah di pahami seperti pada tabel di bawah ini.

Table 1. Kesimpulan hasil dari clustering menggunakan Rapid Miner.

Atribut	Jumlah	Nilai centroid depth
Cluster_0	18,530	57,256
Cluster_1	36,599	45,520
Cluster_2	37,758	48,345

**5. KESIMPULAN DAN SARAN**

Proses clustering menggunakan algoritma K-Means menghasilkan bahwa setiap cluster yang telah dikelompokkan menjadi 3 dengan melihat nilai centroid antara setiap cluster-nya yang menunjukkan gempa dominan. Pada cluste\_ 0 memiliki jumlah paling sedikit yaitu 18,530, namun memiliki nilai centroid kedalaman tertinggi 57,256 yang menunjukkan tipe gempa paling dalam. Pada cluster\_1 memiliki jumlah 36,599 dengan nilai centroid kedalaman 45,520 yang berarti menunjukkan tipe gempa dangkal. Sedangkan pada cluster\_2 menjadi gempa dengan jumlah paling banyak terjadi yaitu 37,758, dan memiliki nilai centroid 48,345 yang berarti gempa dengan kedalaman sedang. Maka yang berbahaya adalah gempa dengan cluster\_1 dengan tipe gempa dangkal yang akan menyebabkan kerusakan hebat. Dengan proses clustering gempa bumi di Indonesia dengan metode K-Means dapat berkontribusi pada penelitian-penelitian

selanjutnya dengan memberikan informasi yang bermanfaat terutama dalam usaha mitigasi dan perencanaan tanggap bencana yang lebih efektif.

#### DAFTAR PUSTAKA

- [1] D. Agung Harijoko , Ade Anggraini, Agung Setianto , Agus Hendratno, Akhmad Solikhin, *Geologi Gempa Bumi Indonesia*. UGM Press, 2023.
- [2] Drs. Sugiharyanto M.Si., *Geografi dan Sosiologi*. Yudhistira Quadra, 2007.
- [3] I. Ernawati, “Data Mining Sebagai Salah Satu Solusi Strategi Bisnis,” *Inform. J. Ilmu Komput.*, vol. 14, no. 1, p. 9, 2018, doi: 10.52958/iftk.v14i1.367.
- [4] I. E. Mustika, Yunita Ardilla, Abraham Manuhutu, Nazaruddin Ahmad, Imanuddin Hasbi, Guntoro, Melda Agnes Manuhutu, Mohamad Ridwan, Hozairi, Anindya Khrisna Wardhani, Syariful Alim, Ikhsan Romli, Yoga Religia, D Tri Octafian, Unggul Utan Sufandi, *DATA MINING DAN APLIKASINYA*. Penerbit Widina, 2021.
- [5] M. T. Muhammad Arhami, S.Si., M.Kom., Muhammad Nasir, S.T., *Data Mining - Algoritma dan Implementasi*. Andi Offset, 2020.
- [6] L. Azzahra and Amru Yasir, “Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 3, no. 1, pp. 1–10, 2024, doi: 10.70340/jirsi.v3i1.88.
- [7] C. P. Ema Ainun Novia, Woro Isti Rahayu, *SISTEM PERBANDINGAN ALGORITMA K-MEANS DAN NAÏVE BAYES UNTUK MEMREDIKSI PRIORITAS PEMBAYARAN TAGIHAN RUMAH SAKIT BERDASARKAN TINGKAT KEPENTINGAN*. Kreatif Industri Nusantara, 2020.
- [8] M. S. Iskandar and Z. Fatah, “Gudang Jurnal Multidisiplin Ilmu Implementasi Metode Algoritma K-Means Clustering Untuk Menentukan Penerima Program Indonesia Pintar ( PIP ),” vol. 2, no. November, pp. 1–8, 2024.
- [9] A. Kusmiran, Minarti, M. F. I. Massinai, A. Zarkasi, A. A. Maharani, and R. Desiani, “Klasifikasi Kedalaman Kejadian Gempa Menggunakan Algoritma K-Means Clustering: Studi Kasus Kejadian Gempa Di Sulawesi,” *JFT J. Fis. dan Ter.*, vol. 9, no. 2, pp. 79–88, 2022, doi: 10.24252/jft.v9i2.29198.
- [10] P. A. Rahmi, T. Muliawati, and D. G. Harbowo, “Analisis Pola Aktivitas Gempa Bumi di Pulau Sumatera dengan Metode K-Means Clustering dan Rantai Markov,” vol. 2024, no. Senada, pp. 592–601, 2024.
- [11] N. Dwitiyanti, S. Ayu Kumala, and S. Dwi Handayani, “PENERAPAN METODE K-MEANS PADA KLASTERISASI WILAYAH RAWAN GEMPA DI INDONESIA Implementation of K-Means Method in Classterization of Earthquake Prone Areas in Indonesia,” *Pros. Semin. Nas. UNIMUS*, vol. 6, pp. 1029–1037, 2023.
- [12] A. W. Zunan Setiawan, Muhammad Fajar, Arif Mudi Priyatno, Anggi Yhurinda Perdana Putri, Mediana Aryuni, Siti Yuliyanti, Harya Widiputra, Budanis Dwi Meilani, Rohmat Nur Ibrahim, Rezania Agramanisti Azdy, Satrio Junaidi, *BUKU AJAR DATA MINING*. 2023.
- [13] N. D. A. Dudih Gustian, Samsul Pahmi, *Menggali Emas Terpendam Data Mining*. 3M Media Karya Serang, 2020.
- [14] L. N. S. Dody Dharma Hutabarat, Canrakerta, Lazuardi Zulfikar Wicaksana, Dimas Rahadian, *Membangun Budaya Data di Kementerian Keuangan*. Central Transformation Office, Sekretariat Jenderal, Kementerian Keuangan, 2021.