

Analisis Prediksi Gempa Bumi di Nusa Tenggara Timur dengan *Metode Naïve Bayes dan K-Means Clustering*

Rosina Senista Tiwe Rani¹, Kukuh Yusdhistiro²
Sistem Informasi, Universitas Merdeka Malang. 081239203409
e-mail: ranirosina@gmail.com¹, k2y0f2t1i@unmer.ac.id²

ABSTRAK

Kata Kunci:

Gempa Bumi
Data Mining
Naïve Bayes
K-Means Clustering

Gempa bumi sulit diprediksi baik dari segi waktu maupun intensitasnya, karena terjadi akibat pelepasan energi di litosfer. Indonesia, yang terletak di Ring of Fire dan persimpangan tiga lempeng utama, sering mengalami gempa tektonik dan vulkanik yang mengakibatkan kerusakan dan korban jiwa. Saat ini, belum ada metode yang akurat untuk memprediksi gempa. Penelitian ini menggunakan data gempa dari Nusa Tenggara Timur dan menerapkan teknik data mining, yaitu *Naïve Bayes* dan *K-Means Clustering*. *Naïve Bayes* menunjukkan akurasi tinggi, terutama untuk tahun 2022, dengan nilai prediksi 0,907 dan *confusion matrix* 90,74%, menandakan tahun tersebut sebagai periode dengan jumlah gempa tertinggi. *K-Means Clustering* mengidentifikasi pola jumlah gempa per tahun, seperti 67 gempa ringan di C1, 7 gempa besar di C2, dan 5 gempa ringan di C3 pada 2022. Hasil ini menunjukkan bahwa metode *Naïve Bayes* dan *K-Means Clustering* efektif untuk memprediksi dan menganalisis gempa, serta menawarkan dasar untuk pengembangan prediksi dan mitigasi risiko gempa di masa depan.

ABSTRACT

Keyword:

Earthquakes
Data Mining
Naïve Bayes
K-Means Clustering

Earthquakes are difficult to predict in terms of timing and intensity due to the release of energy in the lithosphere. Indonesia, located in the Ring of Fire and at the intersection of three major tectonic plates, frequently experiences both tectonic and volcanic earthquakes that cause damage and loss of life. Currently, there is no accurate method for predicting earthquakes. This study uses earthquake data from East Nusa Tenggara and applies data mining techniques, specifically Naïve Bayes and K-Means Clustering. Naïve Bayes demonstrated high accuracy, particularly for the year 2022, with a prediction value of 0.907 and a confusion matrix of 90.74%, indicating that 2022 was a period with the highest number of earthquakes. K-Means Clustering identified patterns in annual earthquake counts, such as 67 minor earthquakes in C1, 7 major earthquakes in C2, and 5 minor earthquakes in C3 for 2022. These results show that Naïve Bayes and K-Means Clustering are effective for predicting and analyzing earthquakes, providing a foundation for future development in earthquake prediction and risk mitigation.

I. PENDAHULUAN

Gempa bumi adalah bencana alam yang sulit diprediksi baik dari segi waktu maupun magnitudo. Terjadi akibat akumulasi energi dalam litosfer yang kemudian menjalar ke permukaan bumi, menyebabkan kerusakan yang signifikan baik pada lingkungan alami maupun buatan (Mustafa, 2010). Gempa bumi terjadi karena dua sebab, yaitu secara tektonik dan vulkanik. Di Indonesia, yang terletak di antara tiga lempeng besar-Eurasia, Indo-Australia, dan Pasifik-serta berada di kawasan *Ring of Fire*, frekuensi gempa bumi, baik tektonik maupun vulkanik.

Proses terjadinya gempa bumi melibatkan interaksi kompleks di bawah permukaan bumi, sehingga sulit untuk diprediksi secara akurat. Aktivitas kegempaan dapat diam selama ratusan atau bahkan ribuan tahun sebelum tiba-tiba melepaskan energi besar yang menyebabkan kerusakan. Hingga saat ini, belum ada teori yang sepenuhnya dapat memprediksi lokasi, waktu, dan magnitudo gempa bumi dengan tepat, sehingga dampaknya tetap sulit diantisipasi. Dengan adanya data gempa bumi di Nusa Tenggara Timur dan kemajuan teknologi, peneliti dapat memanfaatkan teknik data mining untuk membantu dalam analisis kegempaan. Metode klasifikasi seperti Algoritma Naïve Bayes dan K-Means *Clustering* digunakan untuk memproses data dan mengelompokkan informasi ke dalam kelompok dengan beberapa *cluster*. Naïve Bayes dapat memprediksi kejadian gempa di masa depan berdasarkan data historis, sementara K-Means *Clustering* membantu dalam mengelompokkan data gempa menjadi kelompok yang lebih terstruktur. Pendekatan ini memberikan harapan baru dalam memahami pola kegempaan dan meningkatkan upaya mitigasi.

Berdasarkan uraian singkat diatas, peneliti mencoba melakukan penerapan data mining dengan menggunakan penerapan algoritma *naïve bayes* dan *k-means clustering* untuk pengolahan data bencana gempa bumi di kepulauan Nusa Tenggara Timur pada USGS, sehingga akan didapatkan hasil akurasi prediksi mengenai terjadinya gempa bumi di waktu mendatang dari algoritma *naïve bayes* dan *k-means clustering* terhadap data bencana gempa bumi dan dapat diketahui tingkat kesesuaian algoritma *naïve bayes* dan *k-means clustering* untuk diterapkan dalam melakukan prediksi gempa bumi. Dengan tinjauan pustakanya, penerapan data *mining* untuk menganalisis data bencana gempa bumi di kepulauan maluku pada bmkg menggunakan *naïve bayes* algorithm berdasarkan penelitian dapat diketahui bahwa prediksi bencana gempa tahunan pada tahun 2019-2020 di kepulauan maluku yaitu provinsi maluku utara dengan akurasi kebenaran data prediksi 71,1%, di provinsi maluku utara terjadi bencana gempa bumi di pulau ternate dengan akurasi kebenaran 78,6%, di provinsi maluku terjadi bencana gempa bumi di pulau seram dengan kebenaran data prediksi 92%. Berdasarkan penelitian dapat disimpulkan bahwa algoritma *naïve bayes* dapat digunakan dengan baik untuk memprediksi terjadinya gempa bumi, penerapan data mining pada data gempa bumi terhadap potensi tsunami di Indonesia hasil dan pembahasan dapat ditarik kesimpulan bahwasannya dengan menerapkan algoritma *naïve bayes classifier* (NBC) pada data gempa bumi terhadap potensi tsunami di indonesia bisa mengetahui kemungkinan efek yang

diakibatkan pada gempa bumi. Dengan data yang digunakan efek dihasilkan adalah potensi tsunami. Penerapan data mining untuk analisis data bencana milik BNPB menggunakan algoritma k-means dan *linear regression*, penelitian ini membahas tentang pengelompokan jumlah data bencana dan prediksi data bencana yang akan terjadi 5 tahun kedepan menggunakan teknik data mining. Algoritma data mining yang digunakan adalah k-means untuk *clustering* dan *linear regression* untuk prediksi bencana. Analisis prediksi efek kerusakan gempa dari magnitudo (skala richter) dengan metode ID3 menggunakan aplikasi data mining orange, penulis membuat rancangan klasifikasi data mulai dari data *training set*, tahapan perancangan dan penerapan data mining sehingga nantinya akan dihasilkan suatu pola tertentu yang dapat digunakan untuk mempreiksi energi seismik yang dipancarkan oleh sumber gempa. Masalah pada penelitian ini adalah bagaimana penerapan teknik data mining khususnya metode naïve bayes dan k-means clustering dalam memprediksi gempa bumi, apakah kedua metode tersebut dapat memberikan hasil prediksi yang akurat dan dapat diandalkan, bagaimana perbandingan performa antara metode naïve bayes dan k-means clustering dalam memprediksi gempa bumi di Nusa Tenggara Timur, dan tujuan penelitian adalah menerapkan teknik data mining khususnya metode naïve bayes dan k-means clustering untuk memprediksi gempa bumi di Nusa Tenggara Timur. mengukur akurasi dan keandalan prediksi gempa bumi yang dihasilkan oleh metode naïve bayes dan k-means clustering, melakukan perbandingan kinerja antara metode Naïve Bayes dan K-Means Clustering dalam memprediksi gempa bumi.

II. METODE

Penelitian ini menggunakan metode kuantitatif dengan tujuan untuk menganalisis data gempa bumi di Wilayah Nusa Tenggara Timur (NTT) dari tahun 2014 hingga 2022 dan yang telah diperbarui hingga tahun 2023-2024. Populasi penelitian ini meliputi keseluruhan data gempa bumi di Wilayah NTT. Sampel yang digunakan adalah data gempa bumi dari tahun 2014 hingga 2022 yang sudah diupdate tahun 2023-2024. Data tersebut diambil dari sumber sekunder, yaitu dokumen atau laporan yang telah tersedia dari pihak ketiga, dan bukan langsung dari subjek penelitian.

Data dianalisis menggunakan teknik data mining dengan algoritma Naïve Bayes dan K-Means *Clustering*. Algoritma *Naïve Bayes* akan digunakan untuk klasifikasi berdasarkan probabilitas, sementara *K-Means Clustering* akan digunakan untuk mengelompokkan data ke dalam cluster yang relevan. Evaluasi hasil analisis dilakukan dengan menggunakan *Confusion Matrix* untuk menghitung akurasi klasifikasi, yaitu membandingkan jumlah kasus yang berhasil diidentifikasi dengan total kasus yang diuji. Metode ini diharapkan dapat memberikan wawasan lebih dalam mengenai pola kegempaan dan meningkatkan prediksi serta pemahaman tentang kejadian gempa bumi di wilayah tersebut.

III. HASIL DAN PEMBAHASAN

Pada proses pengumpulan data peneliti menggunakan metode Naïve Bayes dan K-Means Clustering untuk menganalisis data gempa bumi di Nusa Tenggara Timur. Data yang dikumpulkan dari website USGS dan mencakup 5 data yaitu, Time, Magnitudo, Depth, Place, MagType

Time	Magnitudo	Depth	Place	MagType
2014	4.6	51.24	94 km SSE of Waingapu, Indonesia	Gempa Ringan
2014	4.4	54.55	87 km S of Waingapu, Indonesia	Gempa Ringan
2014	4	36.97	87 km SE of Waingapu, Indonesia	Gempa Ringan
2014	4.3	28.29	112 km N of Maumere, Indonesia	Gempa Ringan
2014	4.2	54.48	56 km SW of Waingapu, Indonesia	Gempa Ringan
2014	4.1	57.37	56 km WSW of Waingapu, Indonesia	Gempa Ringan
2014	4.2	18.33	49 km NNW of Maumere, Indonesia	Gempa Ringan
2014	4.2	37.12	43 km NNE of Maumere, Indonesia	Gempa Ringan
2014	4.3	24.59	36 km SE of Kupang, Indonesia	Gempa Ringan
2014	4.5	23.46	65 km NNE of Ruteng, Indonesia	Gempa Ringan
2014	5.5	18.6	80 km E of Maumere, Indonesia	Gempa Besar
2014	4.3	35	42 km ESE of Kupang, Indonesia	Gempa Ringan
2014	4.2	42.6	55 km N of Pante Makasar, Timor Leste	Gempa Ringan
2014	4.4	35	73 km E of Maumere, Indonesia	Gempa Ringan
2014	4	37.86	80 km E of Maumere, Indonesia	Gempa Ringan
2014	4.2	32.66	54 km NNE of Ruteng, Indonesia	Gempa Ringan
2015	4.3	51.08	23 km NNW of Labuan Bajo, Indonesia	Gempa Ringan
2015	4.1	12.44	56 km E of Soe, Indonesia	Gempa Ringan
2015	4	51.45	54 km SE of Soe, Indonesia	Gempa Ringan
2015	4.3	25.41	36 km SE of Kupang, Indonesia	Gempa Ringan
2015	4.3	38.87	55 km SE of Soe, Indonesia	Gempa Ringan
2015	4.1	12.89	10 km NNW of Kefamenanu, Indonesia	Gempa Ringan
2015	5	18.8	104 km ENE of Maumere, Indonesia	Gempa Ringan
2015	4.8	42.15	30 km ENE of Kupang, Indonesia	Gempa Ringan
2015	5.7	10.27	49 km NW of Maubara, Timor	Gempa Besar
2015	4.4	52.2	37 km E of Kupang, Indonesia	Gempa Ringan
2015	4.2	18.85	18 km WNW of Maumere, Indonesia	Gempa Ringan
2015	5.3	53.73	38 km WNW of Pante Makasar, Timor Leste	Gempa Ringan
2015	4.4	42.8	31 km SSE of Kefamenanu, Indonesia	Gempa Ringan
2015	5	20.23	62 km NW of Maubara, Timor Leste	Gempa Ringan
2015	4.4	54.67	102 km SSE of Ende, Indonesia	Gempa Ringan
2015	4.5	28.53	53 km ESE of Kupang, Indonesia	Gempa Ringan
2016	4.3	31.05	53 km SE of Soe, Indonesia	Gempa Ringan
2016	5	14.06	17 km N of Kupang, Indonesia	Gempa Ringan
2016	4.5	34.34	16 km NNE of Kupang, Indonesia	Gempa Ringan

Gambar 1. Pengumpulan Data

Pada tahap pengolahan data fokus pada analisis gempa bumi, yang terjadi setiap tahun dalam berbagai skala, mulai dari ringan hingga besar yang menyebabkan tsunami. Data ini menjadi dasar untuk melakukan prediksi dan analisis lebih lanjut. Target utama penelitian ini adalah mengidentifikasi gempa bumi ringan dan besar. Penentuan target ini krusial dalam proses data mining dengan menggunakan metode *Naïve Bayes* dan *K-Means Clustering* untuk memastikan akurasi data yang diolah. Dengan memilih target yang jelas sehingga analisis dapat dilakukan dengan lebih terstruktur dan akurat.

Time	Magnitudo	Depth	Place	MagType
2014	4.6	51.24	94 km SSE of Waingapu, Indonesia	Gempa Ringan
2014	4.4	54.55	87 km S of Waingapu, Indonesia	Gempa Ringan
2014	4	36.97	87 km SE of Waingapu, Indonesia	Gempa Ringan
2014	4.3	28.29	112 km N of Maumere, Indonesia	Gempa Ringan
2014	4.2	54.48	56 km SW of Waingapu, Indonesia	Gempa Ringan
2014	4.1	57.37	56 km WSW of Waingapu, Indonesia	Gempa Ringan
2014	4.2	18.33	49 km NNW of Maumere, Indonesia	Gempa Ringan
2014	4.2	37.12	43 km NNE of Maumere, Indonesia	Gempa Ringan
2014	4.3	24.59	36 km SE of Kupang, Indonesia	Gempa Ringan
2014	4.5	23.46	65 km NNE of Ruteng, Indonesia	Gempa Ringan
2014	5.5	18.6	80 km E of Maumere, Indonesia	Gempa Besar
2014	4.3	35	42 km ESE of Kupang, Indonesia	Gempa Ringan
2014	4.2	42.6	55 km N of Pante Makasar, Timor Leste	Gempa Ringan
2014	4.4	35	73 km E of Maumere, Indonesia	Gempa Ringan
2014	4	37.86	80 km E of Maumere, Indonesia	Gempa Ringan
2014	4.2	32.66	54 km NNE of Ruteng, Indonesia	Gempa Ringan
2015	4.3	51.08	23 km NNW of Labuan Bajo, Indonesia	Gempa Ringan
2015	4.1	12.44	56 km E of Soe, Indonesia	Gempa Ringan
2015	4	51.45	54 km SE of Soe, Indonesia	Gempa Ringan
2015	4.3	25.41	36 km SE of Kupang, Indonesia	Gempa Ringan
2015	4.3	38.87	55 km SE of Soe, Indonesia	Gempa Ringan
2015	4.1	12.89	10 km NNW of Kefamenanu, Indonesia	Gempa Ringan
2015	5	18.8	104 km ENE of Maumere, Indonesia	Gempa Ringan
2015	4.8	42.15	30 km ENE of Kupang, Indonesia	Gempa Ringan
2015	5.7	10.27	49 km NW of Maubara, Timor	Gempa Besar
2015	4.4	52.2	37 km E of Kupang, Indonesia	Gempa Ringan
2015	4.2	18.85	18 km WNW of Maumere, Indonesia	Gempa Ringan
2015	5.3	53.73	38 km WNW of Pante Makasar, Timor Leste	Gempa Ringan
2015	4.4	42.8	31 km SSE of Kefamenanu, Indonesia	Gempa Ringan
2015	5	20.23	62 km NW of Maubara, Timor Leste	Gempa Ringan
2015	4.4	54.67	102 km SSE of Ende, Indonesia	Gempa Ringan
2015	4.5	28.53	53 km ESE of Kupang, Indonesia	Gempa Ringan
2016	4.3	31.05	53 km SE of Soe, Indonesia	Gempa Ringan
2016	5	14.06	17 km N of Kupang, Indonesia	Gempa Ringan
2016	4.5	34.34	16 km NNE of Kupang, Indonesia	Gempa Ringan

Gambar 2. Pengolahan Data dan Penentuan Target

$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + \dots + (X_{ki} - X_{kj})^2}$	(1)
$P(H, X) = P(X \setminus T) (P(T_0 / P(X)))$	(2)

Proses Implementasi Aplikasi Orange

Tahap awal proses pengolahan data adalah menginput database dari excel ke aplikasi orange. Pada tampilan awal orange memiliki beberapa menu diantaranya ada File, CSV File Import, Datasets, SQL Tabel, Data Tabel, Paint Data, Data Info, Rank, Edit Domain, Color, Feature Statistics, Save Data dengan fungsi sebagai penyelesaian masalah yang terdapat pada data mining.

Proses memilih dataset yang akan diinput, dataset yang digunakan merupakan Data Gempa Bumi di Nusa Tenggara Timur 11 tahun terakhir yaitu tahun 2014-2024. Dataset akan diolah menggunakan aplikasi orange, dengan target yang sudah ditentukan yaitu gempa ringan dan gempa besar. Pada data gempa bumi ini fitnya terdiri dari Time, Mag, Depth, Place, MagType.

Hasil prediksi Naïve Bayes untuk keseluruhan Data Gempa Bumi

Setelah melakukan proses data mining maka bisa menghasilkan nilai akurasi pada setiap tahun yang berbeda-beda yang dilihat dari dataset yang telah di input pada proses data mining. Prediksi yang telah ditarget pada setiap tahun adalah Gempa Ringan dan Gempa Besar. Data yang terdiri dari 272 kolom dengan hasil model naïve bayes.

AUC (Area Under the Curve) = 0.934

CA (Classification Accuracy) = 0.898

F1= 0.897

Prec= 0.897

Recall = 0.898

MCC = 0.504

Berdasarkan metode Naïve Bayes CA merupakan hasil akurasi dataset Gempa Bumi untuk keseluruhan dengan nilai akurasinya 0.898.

	MagType	Place	Time	Magnitudo	Depth
1	Gempa Ringan	94 km SSE of Waingapu, Indonesia	2014	4.6	51.240
2	Gempa Ringan	87 km S of Waingapu, Indonesia	2014	4.4	54.550
3	Gempa Ringan	87 km SE of Waingapu, Indonesia	2014	4.0	36.970
4	Gempa Ringan	112 km N of Maumere, Indonesia	2014	4.3	28.290
5	Gempa Ringan	56 km SW of Waingapu, Indonesia	2014	4.2	54.480
6	Gempa Ringan	56 km WSW of Waingapu, Indonesia	2014	4.1	57.370
7	Gempa Ringan	49 km NNW of Maumere, Indonesia	2014	4.2	18.330
8	Gempa Ringan	43 km NNE of Maumere, Indonesia	2014	4.2	37.120
9	Gempa Ringan	36 km SE of Kupang, Indonesia	2014	4.3	24.590
10	Gempa Ringan	65 km NNE of Ruteng, Indonesia	2014	4.5	23.460
11	Gempa Besar	80 km E of Maumere, Indonesia	2014	5.5	18.600
12	Gempa Ringan	42 km ESE of Kupang, Indonesia	2014	4.3	35.000
13	Gempa Ringan	55 km N of Pante Makasar, Timor Leste	2014	4.2	42.600
14	Gempa Ringan	73 km E of Maumere, Indonesia	2014	4.4	35.000
15	Gempa Ringan	80 km E of Maumere, Indonesia	2014	4.0	37.860
16	Gempa Ringan	54 km NNE of Ruteng, Indonesia	2014	4.2	32.660
17	Gempa Ringan	23 km NNW of Labuan Bajo, Indonesia	2015	4.3	51.080
18	Gempa Ringan	56 km E of Soe, Indonesia	2015	4.1	12.440
19	Gempa Ringan	54 km SE of Soe, Indonesia	2015	4.0	51.450
20	Gempa Ringan	36 km SE of Kupang, Indonesia	2015	4.3	25.410
21	Gempa Ringan	55 km SE of Soe, Indonesia	2015	4.3	38.870
22	Gempa Ringan	10 km NNW of Kefamenanu, Indonesia	2015	4.1	12.890
23	Gempa Ringan	104 km ENE of Maumere, Indonesia	2015	5.0	18.800
24	Gempa Ringan	30 km ENE of Kupang, Indonesia	2015	4.8	42.150
25	Gempa Besar	49 km NW of Maubara, Timor	2015	5.7	10.270
26	Gempa Ringan	37 km E of Kupang, Indonesia	2015	4.4	52.200
27	Gempa Ringan	18 km WNW of Maumere, Indonesia	2015	4.2	18.850
28	Gempa Ringan	38 km WNW of Pante Makasar, Timor Leste	2015	5.3	53.730
29	Gempa Ringan	31 km SSE of Kefamenanu, Indonesia	2015	4.4	42.800
30	Gempa Ringan	62 km NW of Maubara, Timor Leste	2015	5.0	20.230
31	Gempa Ringan	102 km SSE of Ende, Indonesia	2015	4.4	54.670
32	Gempa Ringan	53 km ESE of Kupang, Indonesia	2015	4.5	28.530
33	Gempa Ringan	53 km SE of Soe, Indonesia	2016	4.3	31.050
34	Gempa Ringan	17 km N of Kupang, Indonesia	2016	5.0	14.060
35	Gempa Ringan	16 km NNE of Kupang, Indonesia	2016	4.5	34.340
36	Gempa Ringan	Timor region, Indonesia	2016	5.2	10.000
37	Gempa Ringan	136 km SSE of Ende, Indonesia	2016	4.3	45.840
38	Gempa Ringan	115 km SSE of Ende, Indonesia	2016	4.0	54.530
39	Gempa Besar	39 km W of Waingapu, Indonesia	2016	5.4	48.880
40	Gempa Ringan	89 km WNW of Maubara, Timor Leste	2016	4.6	27.610

Model	AUC	CA	F1	Prec	Recall	MCC
Naive Bayes	0.934	0.898	0.897	0.897	0.898	0.504

Gambar 3. Hasil Prediksi untuk keseluruhan Data Gempa Bumi

Hasil Confusion Matrix untuk Keseluruhan Data Gempa Bumi

Confusion Matrix adalah pengukuran performa dalam permasalahan klasifikasi yang dimana berbentuk dua kelas. Confusion Matrix memiliki 4 tabel dengan arti yang berbeda-

beda dari hasil tabel prediksi atau prediksi serta nilai yang aktual. *Confusion Matrix* juga merupakan tabel yang memberikan tampilan hasil dari suatu kerja berdasarkan suatu klasifikasi dari data set mining. Dengan menghitung akurasi yang paling penting pada suatu dataset prediksi Gempa Bumi.

1. *True Positif* (TP) prediksi gempa besar yang benar
2. *True Negatif* (TN) prediksi gempa ringan yang benar
3. *False Positif* (FP) prediksi gempa besar yang salah
4. *False Negatif* (FN) prediksi gempa ringan yang salah

		Predicted		Σ
		Gempa Besar	Gempa Ringan	
Actual	Gempa Besar	61	49	110
	Gempa Ringan	46	774	820
Σ		107	823	930

Gambar 4. Hasil *Confusion Matrix* untuk Keseluruhan Data Gempa Bumi

Confusion Matrix untuk keseluruhan data gempa bumi pada gambar diatas menunjukkan nilai *True Positif* (TP) 774 *True Negatif* (TN) 61, *False Positif* (FP) 49, dan *False Negatif* (FN) 46.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\%$$

$$\text{Accuracy} = \frac{(774+61)}{(774+61+49+46)} \times 100\%$$

$$\text{Accuracynya} = 89.79\%$$

$$\text{Precision} = \frac{(TP)}{(TP+FP)} \times 100\%$$

$$\text{Precision} = \frac{(774)}{(774 + 49)} \times 100\%$$

$$\text{Precision} = 94.07\%$$

$$\text{Recall} = \frac{(TP)}{(TP+FN)} \times 100\%$$

$$\text{Recall} = \frac{(774)}{774+46} \times 100\%$$

$$\text{Recall} = 94.39\%$$

Berdasarkan perhitungan akurasi pada *tools Orange Prediksi dan Confusion Matrix* Gempa Bumi di Nusa Tenggara Timur selama 11 tahun (2014-2024)

Tabel 1. Hasil Prediksi

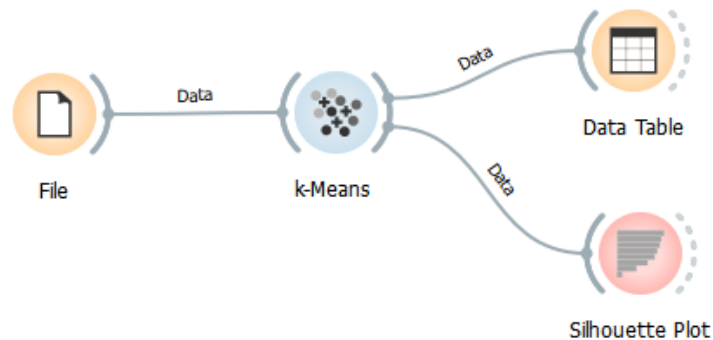
Tahun	AUC	CA	F1	Prediksi	Recall	MCC
2014	0.918	0.890	0.890	0.890	0.890	0.470
2015	0.278	0.800	0.800	0.800	0.800	-0.111
2016	0.997	0.812	0.850	0.940	0.812	0.503
2017	0.174	0.760	0.795	0.832	0.760	-0.129
2018	0.824	0.817	0.812	0.848	0.817	0.664
2019	0.478	0.800	0.809	0.820	0.800	0.110
2020	0.657	0.633	0.679	0.782	0.633	0.183
2021	0.590	0.850	0.850	0.850	0.850	0.018
2022	0.908	0.907	0.895	0.885	0.907	0.156
2023	0.828	0.790	0.803	0.829	0.790	0.452
2024	0.828	0.790	0.803	0.829	0.790	0.452

Tabel 2. Hasil *Confusion Matrix*

Tahun	Accuracy	Precision	Recall
2014	75%	95.74%	77.59%
2015	80%	88.89%	88.89%
2016	81.25%	100%	79.45%
2017	76%	90.48%	82.61%
2018	81.67%	95.24%	66.67%
2019	80%	90%	87.10%
2020	63.33%	88.89%	64%
2021	85%	91.82%	91.82%
2022	90.74%	93.44%	96.8%
2023	79%	91.55%	81.25%
2024	71.67%	88.37%	76%

Clustering Menggunakan Metode K-Means

Widget yang digunakan pada metode K-Means dalam penelitian ini yaitu File, K-Means, Data Tabel, Silhouette Plot. Widget File diambil dari kelompok Data, K-Means diambil dari Unsupervised, Data Tabel diambil dari Data dan Silhouette diambil dari Visualize. Widget merupakan elemen kecil tambahan yang dapat ditampilkan dan dipasang di dalam halaman sebuah web atau aplikasi tertentu. Berikut merupakan gambar dari metode K-Means.

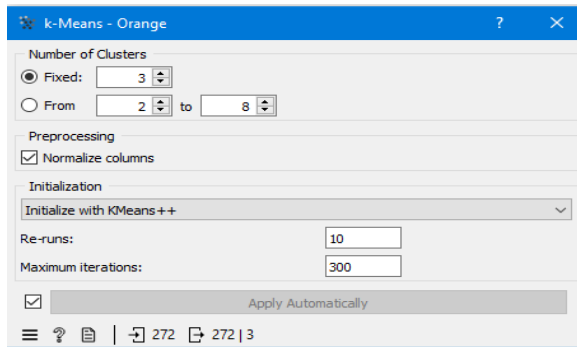


Gambar 5. Tampilan untuk Metode K-Means

Clustering K-Means untuk keseluruhan Data Gempa Bumi

1. Penentuan Cluster Untuk Keseluruhan Data Gempa Bumi

Pada tahap ini penentuan cluster yang akan digunakan sesuai dengan tabel *clustering* yang telah ditentukan yaitu mengenai Gempa Ringan dan Gempa Besar. Pada proses pengaturan K-Means di Orange maka ditentukan Fixed 3 untuk melihat hasil dari setiap cluster.



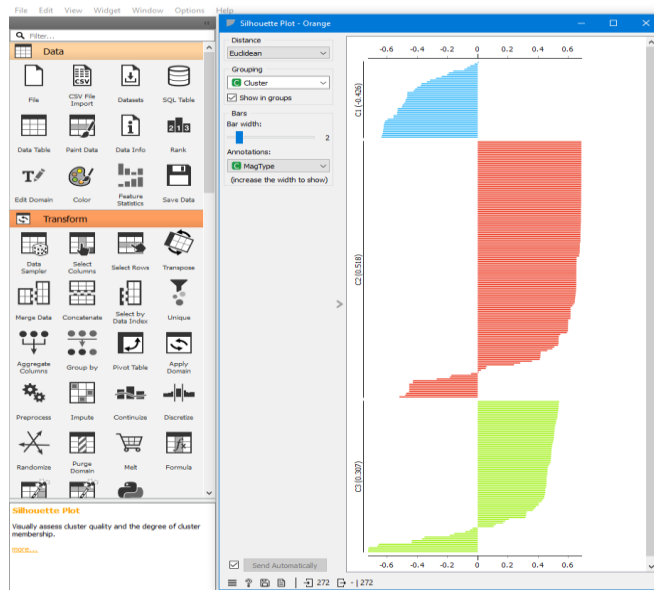
Gambar 6. Penentuan Proses Cluster Keseluruhan Data Gempa Bumi

2. Hasil Clustering Data Tabel

Data tabel merupakan *widget* untuk menampilkan data dalam bentuk *spreadsheet*. Dalam data tabel ini untuk melihat hasil pengelompokannya. Dapat dilihat jika pada tabe C1 dengan jumlah 43 kali yang terkena gempa besar dan gempa ringan, pada C2 terdapat 144 kali yang terkena gempa ringan, pada C3 terdapat 85 kali yang mengalami gempa ringan.

	Place	Cluster	Silhouette	Time	Magnitudo	Depth	MagType
1	94 km SSE of Waingapu, Indonesia	C3	0.640722	2014	4.6	51.240	Gempa Ringan
2	87 km S of Waingapu, Indonesia	C3	0.646829	2014	4.4	54.550	Gempa Ringan
3	87 km SE of Waingapu, Indonesia	C3	0.640258	2014	4.0	36.970	Gempa Ringan
4	112 km N of Maumere, Indonesia	C3	0.628754	2014	4.3	28.290	Gempa Ringan
5	56 km SW of Waingapu, Indonesia	C3	0.647446	2014	4.2	54.480	Gempa Ringan
6	56 km WSW of Waingapu, Indonesia	C3	0.642956	2014	4.1	57.370	Gempa Ringan
7	49 km NNW of Maumere, Indonesia	C3	0.587443	2014	4.2	18.330	Gempa Ringan
8	43 km NNE of Maumere, Indonesia	C3	0.647292	2014	4.2	37.120	Gempa Ringan
9	36 km SE of Kupang, Indonesia	C3	0.615514	2014	4.3	24.590	Gempa Ringan
10	65 km NNE of Ruteng, Indonesia	C3	0.604676	2014	4.5	23.460	Gempa Ringan
11	80 km E of Maumere, Indonesia	C1	0.594265	2014	5.5	18.600	Gempa Besar
12	42 km ESE of Kupang, Indonesia	C3	0.644935	2014	4.3	35.000	Gempa Ringan
13	55 km N of Pante Makasar, Timor Leste	C3	0.651363	2014	4.2	42.600	Gempa Ringan
14	73 km E of Maumere, Indonesia	C3	0.643375	2014	4.4	35.000	Gempa Ringan
15	80 km E of Maumere, Indonesia	C3	0.641375	2014	4.0	37.860	Gempa Ringan
16	54 km NNE of Ruteng, Indonesia	C3	0.639725	2014	4.2	32.660	Gempa Ringan
17	23 km NNW of Labuan Bajo, Indonesia	C3	0.653975	2015	4.3	51.080	Gempa Ringan
18	56 km E of Soe, Indonesia	C3	0.542367	2015	4.1	12.440	Gempa Ringan
19	54 km SE of Soe, Indonesia	C3	0.647302	2015	4.0	51.450	Gempa Ringan
20	36 km SE of Kupang, Indonesia	C3	0.616597	2015	4.3	25.410	Gempa Ringan
21	55 km SE of Soe, Indonesia	C3	0.653416	2015	4.3	38.870	Gempa Ringan
22	10 km NNW of Kefamenanu, Indonesia	C3	0.545168	2015	4.1	12.890	Gempa Ringan
23	104 km ENE of Maumere, Indonesia	C1	0.49188	2015	5.0	18.800	Gempa Ringan
24	30 km ENE of Kupang, Indonesia	C3	0.622878	2015	4.8	42.150	Gempa Ringan
25	49 km NW of Maubara, Timor	C1	0.613254	2015	5.7	10.270	Gempa Besar
26	37 km E of Kupang, Indonesia	C3	0.651995	2015	4.4	52.200	Gempa Ringan
27	18 km WNW of Maumere, Indonesia	C3	0.5816	2015	4.2	18.850	Gempa Ringan
28	38 km WNW of Pante Makasar, Timor Leste	C3	0.545993	2015	5.3	53.730	Gempa Ringan
29	31 km SSE of Kefamenanu, Indonesia	C3	0.654754	2015	4.4	42.800	Gempa Ringan
30	62 km NW of Maubara, Timor Leste	C1	0.484969	2015	5.0	20.230	Gempa Ringan
31	102 km SSE of Ende, Indonesia	C3	0.649558	2015	4.4	54.670	Gempa Ringan
32	53 km ESE of Kupang, Indonesia	C3	0.623469	2015	4.5	28.530	Gempa Ringan
33	53 km SE of Soe, Indonesia	C3	0.633518	2016	4.3	31.050	Gempa Ringan
34	17 km N of Kupang, Indonesia	C1	0.515031	2016	5.0	14.060	Gempa Ringan
35	16 km NNE of Kupang, Indonesia	C3	0.638067	2016	4.5	34.340	Gempa Ringan
36	Timor region, Indonesia	C1	0.5455	2016	5.2	10.000	Gempa Ringan
37	136 km SSE of Ende, Indonesia	C3	0.655989	2016	4.3	45.840	Gempa Ringan
38	115 km SSE of Ende, Indonesia	C3	0.644815	2016	4.0	54.530	Gempa Ringan
39	39 km W of Waingapu, Indonesia	C1	0.532833	2016	5.4	48.880	Gempa Besar
40	89 km WNW of Maubara, Timor Leste	C3	0.603921	2016	4.6	27.610	Gempa Ringan
41	51 km SE of Kupang, Indonesia	C3	0.483302	2016	4.4	10.000	Gempa Ringan
42	32 km SE of Soe, Indonesia	C3	0.618296	2016	4.5	28.960	Gempa Ringan

Gambar 7. Hasil Clustering Data Tabel Keseluruhan Gempa Bumi



Gambar 8. Silhouette Plot Keseluruhan Data Gempa Bumi

Hasil *Clustering* menggunakan metode K-Means dalam penelitian ini menunjukkan jumlah gempa bumi ringan dan besar setiap tahun. Misalnya, pada tahun 2022 terdapat 79 kolom dengan hasil clusternya yaitu C1 67 kolom yang mengalami gempa ringan, C2 7 kolom gempa besar, dan C3 5 kolom yang mengalami gempa ringan.

IV. SIMPULAN

Tujuan utama dalam penelitian ini adalah menerapkan teknik data mining dengan menggunakan metode Naïve Bayes dan K-Means Clustering dalam memprediksi gempa bumi di Nusa Tenggara Timur, serta mengukur akurasi, keandalan, dan melakukan perbandingan kinerja antara kedua metode, untuk melakukan prediksi jumlah atau klasifikasi kejadian gempa bumi metode yang lebih akurat dan cocok untuk masalah ini adalah dengan menggunakan metode Naïve Bayes. Meskipun Naïve Bayes memberikan prediksi yang baik, penggunaan K-Means Clustering untuk analisis pola gempa bumi juga memberikan wawasan tambahan tentang distribusi dan karakteristik gempa bumi setiap tahun. Hasil perhitungan menggunakan metode Naïve Bayes, Tahun 2022 menunjukkan hasil prediksi terbaik dengan nilai 0.907 dan confusion matrix 90.74%, penelitian ini menunjukkan bahwa tahun 2022 mengalami jumlah gempa ringan dan besar terbanyak dalam 11 tahun terakhir. Hasil Clustering menggunakan metode K-Means dalam penelitian ini menunjukkan jumlah gempa bumi ringan dan besar setiap tahun. Misalnya, pada tahun 2022 hasil clustering menunjukkan total gempa 79 kolom dengan hasil clusternya yaitu C1 sebanyak 67 kolom yang mengalami gempa ringan, C2 terdapat 7 kolom dengan gempa besar, dan C3 sebanyak 5 kolom yang mengalami gempa ringan.

DAFTAR RUJUKAN

- [1] Althaf, M. A. P. P. (2023). Analisis Probabilitas Bencana Alam dengan Penerapan Data Mining Menggunakan K-Means dan Linier Regression. *Jurnal Mahasiswa Teknik Informatika*, 2(2), 1-7.
- [2] Dari, A. P. E. K. G. Magnitudo (Skala Richter) Dengan Metode Algoritma Id3 Menggunakan Aplikasi Data Mining Orange.
- [3] EMILIA, D. (2021). GEMPA BUMI (Bencana Alam yang Sulit Diprediksi) - Institutional Repository of UIN SATU Tulungagung. [Uinsatu.ac.id. http://repo.uinsatu.ac.id/40984/1/COVER.pdf](http://repo.uinsatu.ac.id/40984/1/COVER.pdf)
- [4] Handoko, K. (2016). Penerapan Data Mining dalam Meningkatkan Mutu Pembelajaran Menggunakan Metode K-MEANS Clustering. *Jurnal Nasional Teknologi dan Sistem Informasi*, 2(3), 31-40.
- [5] Harahap, C. N., Reviantika, F., & Azhar, Y. (2020). Analisis Gempa Bumi Pada Pulau Jawa Menggunakan Clustering Algoritma K-Means. *Jurnal Dinamika Informatika*, 9(1), 51-60.
- [6] Ismail, I. (2021). Klasifikasi Area Gempa Bumi Menggunakan Algoritma Random Forest. *Jurnal Ilmiah Informatika Komputer*, 26(1), 56-64.
- [7] Laia, F., 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.
- [8] Lumban Raja, A. E. (2023). Implementasi Data Mining dengan Algoritma Naive Bayes untuk Klasifikasi Kelayakan Penerima Bantuan Sembako (Doctoral dissertation, Prodi Teknik Informatika).
- [9] Mustafa, B. (2010, March 30). Analisis Gempa Nias Dan Gempa Sumatera Barat Dan Kesamaannya Yang Tidak Menimbulkan Tsunami. [ResearchGate; Universitas Andalas. https://www.researchgate.net/publication/325201727_ANALISIS_GEMPA_NIAS_DAN_GEMPA_SUMATERA_BARAT_DAN_KESAMAANNYA_YANG_TIDAK_MENIMBULKAN_TSUNAMI](https://www.researchgate.net/publication/325201727_ANALISIS_GEMPA_NIAS_DAN_GEMPA_SUMATERA_BARAT_DAN_KESAMAANNYA_YANG_TIDAK_MENIMBULKAN_TSUNAMI)
- [10] Ramadhan, M. I., & Prihandoko, P. (2019). Penerapan Data Mining Untuk Analisis Data Bencana Milik BNPB Menggunakan Algoritma K-Means dan Linear Regression. *Jurnal Ilmiah Informatika Komputer*, 22(1).
- [11] Rizal, S., & qusnul Khotimah, R. (2022). Penerapan Data Mining Untuk Clustering Data Penduduk Yang Terdampak Covid-19 Menggunakan Algoritma K-Means. *Jurnal Pendidikan dan Konseling (JPDK)*, 4(4), 2781-2792.
- [12] Surojudin, N. (2020). Penerapan Data Mining Untuk Menganalisis Data Bencana Gempa Bumi Di Kepulauan Maluku Pada BMKG Menggunakan Naïve Bayes Algorithm. *Jurnal SIGMA*, 11(3), 167-174.
- [13] Utomo, D. P., & Purba, B. (2019, September). Penerapan datamining pada data gempa bumi

- terhadap potensi tsunami di Indonesia. In Prosiding Seminar Nasional Riset Information Science (SENARIS) (Vol. 1, pp. 846-853).
- [14] Wahyu, A., & Rushendra, R. (2022). Klasterisasi Dampak Bencana Gempa Bumi Menggunakan Algoritma K-Means di Pulau Jawa. JEPIN (Jurnal Edukasi dan Penelitian Informatika), 8(1), 174-179.
- [15] Zubair, A. (2022). Data Mining Menggunakan Orange. Yogyakarta: Teknosain.