



PREDIKSI DAMPAK GEMPA BUMI DI INDONESIA DENGAN MENGGUNAKAN *ARTIFICIAL NEURAL NETWORK*

Kevin Sidharta Handoyo¹, Yosefina Finsensia Riti², Ricky Junianto Wijaya³

¹kevin.handoyo@student.ukdc.ac.id, ²yosefina.riti@ukdc.ac.id, ³ricky.wijaya@student.ukdc.ac.id
^{1,2,3}Ilmu Informatika, Universitas Katolik Darma Cendika

Abstrak

Penelitian ini memiliki tujuan untuk membangun sistem prediksi dampak gempa bumi di Indonesia dengan mengimplementasikan metode kecerdasan buatan yaitu *Artificial Neural Network* (ANN). Dataset yang digunakan adalah data gempa bumi Indonesia tahun 2023–2025 yang telah melalui proses *preprocessing*, *labeling*, dan penyeimbangan kelas dengan metode SMOTE. Model ANN dirancang dengan 24 *input neuron*, dua *hidden layer*, dan *output* 3 kelas dampak. Pelatihan model dilakukan menggunakan data *training* sebesar 70% dan data *testing* sebesar 30%. Evaluasi dilakukan dengan menggunakan matrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil yang diperoleh menunjukkan *accuracy* ANN sebesar 99% dan *F1-score* tinggi pada semua kelas termasuk kelas minoritas. Pengujian lanjutan menunjukkan model tetap akurat dalam memprediksi dampak berdasarkan input magnitudo dan kedalaman.

Kata kunci: Gempa Bumi, Prediksi Dampak, *Artificial Neural Network*, SMOTE, Klasifikasi.

Abstract

This study aims to develop an earthquake impact prediction system in Indonesia by implementing the an artificial intelligence method, namely ANN. The dataset used consists of Indonesian earthquake data from a 2023–2025, which has undergone preprocessing, labeling, and class balancing using the SMOTE method. The ANN model is designed with 24 input neurons, two hidden layers, and an output layer representing three impact classes. The model was trained using 70% of the data and tested with the remaining 30%. Evaluation was carried out using performance metrics such as accuracy, precision, recall, and F1-score. The results showed that the ANN achieved an accuracy of 99% and high F1-scores across all classes, including the minority class. Further testing confirmed that the model remained accurate in predicting earthquake impacts based on input magnitude and depth.

Keywords: Earthquake, Impact Prediction, *Artificial Neural Network*, SMOTE, Classification.

1. Pendahuluan

Indonesia merupakan sebuah Negara yang terdiri dari ribuan pulau yang berada pada pertemuan tiga lempeng tektonik yang aktif dan saling bergerak satu sama lain, keadaan ini membuat Indonesia sebagai wilayah yang rawan akan bencana gempa bumi [1]. Gempa bumi yang menyebabkan hilangnya nyawa, kerusakan struktur bangunan, lingkungan alam, serta menimbulkan kerugian harta benda disebut dengan *Destructive Earthquake* atau gempa bumi merusak [2]. Di Indonesia beberapa kali terjadi gempa bumi merusak, seperti pada tahun 1856 dengan skala sebesar VI-VII dan pada 2014 sebesar IV-V [3]. Menurut Badan Meteorologi, Klimatologi, dan Geofisika atau BMKG, kejadian gempa yang terjadi pada tahun 2024 di Bandung juga merupakan kategori merusak dan paling tinggi dalam periode 24 tahun terakhir, gempa tersebut telah menghancurkan setidaknya 2.020 unit rumah di Kabupaten Bandung dan 204 unit rumah di Kabupaten Garut dengan tercatat 68 orang luka dan 1 orang meninggal dunia [4].

Gempa bumi sering kali tidak dapat diprediksi secara tepat waktu atau akurat, dikarenakan kompleksitas geologi dan dinamika dari kerak bumi di wilayah tersebut [5]. Para ahli geografi belum menemukan teori yang sangat akurat untuk memprediksi waktu dan lokasi gempa bumi di Indonesia, meskipun bisa mengindikasikan beberapa wilayah akan mengalami gempa besar namun dampaknya masih sulit untuk diprediksi [6]. Prediksi dampak gempa bumi sangat diperlukan oleh berbagai pihak khususnya pemerintah dalam upaya pencegahan risiko bencana yang bisa digunakan sebagai dasar kebijakan dalam menanggulangi gempa bumi [7].

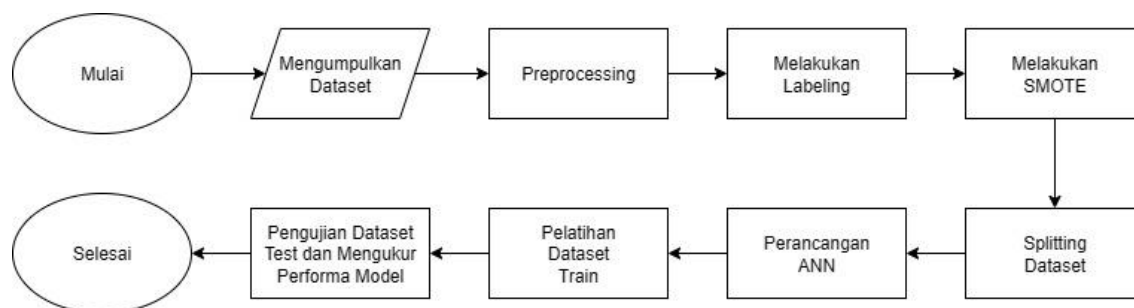
Berbagai model telah digunakan untuk prediksi gempa bumi, seperti *Support Vector Machine* (SVM) dengan hasil menunjukkan nilai *Root Mean Square Error* (RMSE) terbaik sebesar 0,712 [7]. *Naive Bayes* juga digunakan untuk memprediksi gempa bumi di Yogyakarta berdasarkan nilai *magnitudo* dengan hasil akurasi 90,91%, berdasarkan nilai kedalaman dengan hasil akurasi 81,82%, dan berdasarkan lokasi dengan hasil akurasi sebesar 90,91% [6]. Model *Artificial Neural Network* (ANN) yang digunakan untuk memprediksi gempa, diperoleh hasil nilai error dari arsitektur terbaik yaitu sebesar 0,1958 menggunakan 9 *node* atau data dari 9 bulan [8].

SVM memiliki kelebihan dalam mengklasifikasi suatu dataset sebagai titik dalam ruang vektor secara *non-linear*, namun SVM sulit dalam pemilihan parameter dengan optimal saat berhadapan dengan *dataset* yang besar dan kompleks [9]. *Naive Bayes* memiliki kelebihan seperti kemampuan menyesuaikan jumlah dan titik pada data yang berbeda serta melakukan prediksi nilai peluang, namun memiliki kelemahan pada asumsi antar fitur yang menyebabkan ketidaksesuaian pada kondisi aslinya, hal ini berpengaruh pada menurunnya akurasi di beberapa situasi [10]. Sedangkan ANN sendiri memiliki keunggulan dalam merepresentasikan model baik secara *linier* juga *non-linier* dengan jangkauan yang cukup luas dan mampu mendeteksi secara akurat dan komplit tanpa keraguan walaupun dalam hubungan *non-linear* [11]. Kelebihan ANN ini tepat untuk karakteristik *dataset* gempa bumi yang besar dan kompleks, namun jumlah titik atau *node* berpengaruh signifikan terhadap nilai *error*. Oleh karena itu, ANN dipilih untuk memprediksi gempa bumi di Indonesia, namun dengan meningkatkan jumlah *node* dari *dataset*.

Pada penelitian ini, dilakukan peningkatan penggunaan jumlah *node* atau jumlah data pada model ANN yang digunakan, sehingga hasil dari prediksi gempa bumi yang diperoleh menjadi lebih akurat dan diperoleh nilai *error* yang lebih baik.

2. Metode

Pada metode penelitian terdapat beberapa tahapan untuk menyelesaikan masalah. Tahapan pertama yaitu mengumpulkan *dataset*, di tahap ini akan dilakukan pengumpulan *dataset* gempa bumi di Indonesia dari tahun 2023 sampai 2025. Di tahap kedua dilakukan *preprocessing* dengan melakukan data *cleaning* dan normalisasi. Pada tahap ketiga, dilakukan proses *labeling* sebagai *output* untuk model. Di tahap keempat dilakukan penambahan jumlah data kelas minoritas untuk menghindari *overfitting* terhadap kelas mayoritas. Lalu di tahap kelima dilakukan *splitting dataset* untuk membagi data menjadi data untuk *train* dan untuk *test*. Di tahap keenam dilakukan perancangan arsitektur *Artificial Neural Network* (ANN) dengan menggunakan *activation function*, *optimizer*, dan *loss function*. Pada tahap ketujuh baru dilakukan pelatihan model menggunakan ANN. Setelah itu di tahap kedelapan dilakukan pengujian performa prediksi menggunakan matrik evaluasi. Terlihat pada Gambar 1 merupakan kerangka kerja penelitian yang digunakan pada penyelesaian masalah dalam metode penelitian.



Gambar 1. Kerangka Kerja Penelitian

2.1 Pengumpulan Dataset

Dataset yang dipakai pada penelitian ini merupakan data gempa bumi yang terjadi di Indonesia mulai tahun 2008-2025 dengan total 121.717 data, yang diambil dari *Kaggle* [12] dengan judul *Earthquakes in Indonesia*. Sampel yang dipilih untuk dipakai pada penelitian ini merupakan data dari tahun 2023-2025 dengan total 24 bulan. Pada Gambar 2 dibawah ini merupakan contoh 5 data gempa bumi dari *dataset* yang dipakai pada penelitian ini.

	A	B	C	D	E	F	G	H	I	J
1	eventID	datetime	latitude	longitude	magnitude	mag_type	depth	phasecount	azimuth_gap	location
2	bmg2008vkye	2008-11-01 00:31:25.143741+00:00	-0.60443985	98.895531	2.989742	MLv	20		6	146.70738 Southern Sumatra, Indonesia
3	bmg2008vlaj	2008-11-01 01:34:29.660856+00:00	-6.6118598	129.38722	5.5075487	mb	30	62	45.46764	Banda Sea
4	bmg2008vlaj	2008-11-01 01:38:14.802129+00:00	-3.6505861	127.99068	3.5396736	MLv	5	4	331.97404	Seram, Indonesia
5	bmg2008vlbt	2008-11-01 02:20:05.909515+00:00	-4.1989245	128.097	2.4243136	MLv	5	5	326.37993	Banda Sea

Gambar 2. Dataset Gempa Bumi Indonesia

2.2 Preprocessing

Preprocessing data berguna untuk mengolah data mentah menjadi data yang siap untuk dipakai dan diproses lebih lanjut, tahapan ini dilakukan untuk membersihkan data, menghapus data bermasalah, dan mentransformasikan data [13]. Data yang tidak lengkap akan dihapus dan format data yang terlalu rinci akan disederhanakan untuk menghindari kompleksitas berlebih yang dapat mengganggu fokus model dalam proses *training* [14]. Proses *preprocessing* ini melibatkan beberapa langkah seperti data *cleaning*, seleksi fitur, normalisasi, dan perubahan data kategorikal menjadi numerik agar bisa diolah oleh *machine learning* [15]. Pada Gambar 3 menunjukkan *dataset* sebelum *preprocessing*.

eventID	datetime	latitude	longitude	magnitude	mag_type	depth	phasecount	azimuth_gap	location	...	doubleCouple	ciwd	strikeM1	dipM1	rakeM1	strikeM2	dipM2	rakeM2	azgapM1	misfit
0	bmg2022zrajn 2023-01-01 00:22:07.771	5289787769	956662598	3.659852e+09	M	10	20	1230315323	Northern Sumatra, Indonesia	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	bmg2022zqlu 2023-01-01 01:01:37.065	6917235374	1071124344	1.757682e+09	MLv	26	8	1348132973	Java, Indonesia	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	bmg2022zqpo 2023-01-01 02:55:14.210	2334151983	1271005402	3.739536e+09	M	13	21	1110307046	Northern Maluku Sea	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	bmg2022zqpo 2023-01-01 03:55:52.782	3383604546	9707469177	2.955640e+09	M	34	21	2065423584	Northern Sumatra, Indonesia	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	bmg2022zqj 2023-01-01 05:21:43.961	8506458402	1209354782	2.759180e+09	M	11	15	1643457794	Minahasa Peninsula, Sulawesi	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Gambar 3. Dataset Sebelum *Preprocessing*

2.3 Labeling

Pada tahap ini, dilakukan proses pelabelan dataset untuk mengukur hasil prediksi dari model yang dikembangkan menggunakan metode klasifikasi. Proses pelabelan dilakukan dengan memastikan setiap data dalam dataset diberikan label yang sesuai berdasarkan pemahaman mendalam terkait objek yang diteliti [16].

2.4 Synthetic Minority Oversampling Technique (SMOTE)

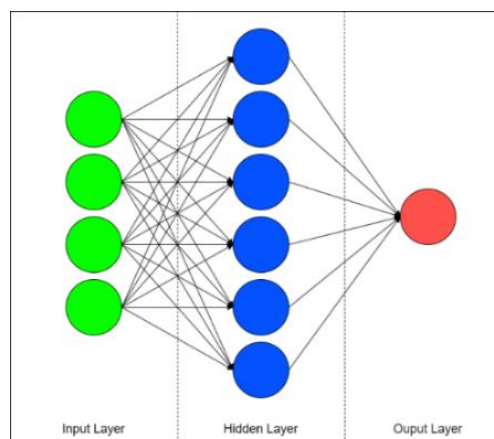
Pada tahap ini, SMOTE digunakan untuk melakukan *oversampling* terhadap kelas minoritas untuk menghasilkan sampel sintesis baru [17]. SMOTE dapat menghasilkan data sintesis baru dari kelas minoritas untuk mengatasi masalah *overfitting*, SMOTE akan melakukan interpolasi linear antara sampel minoritas dan sampel tetangga untuk menghasilkan sampel sintesis [17].

2.5 Splitting Dataset

Split dataset merupakan teknik yang dipakai untuk membagi *dataset*, teknik ini akan mempengaruhi seberapa baik performa dari model yang akan dilatih. Proses *split* dataset akan membagi kumpulan data menjadi dua yaitu data *training* dan data *testing*, yang mana nanti model klasifikasi akan dilatih memakai data *training* dan diuji memanfaatkan data *testing* [18].

2.6 Perancangan Artificial Neural Network

ANN atau yang dikenal juga sebagai *Artificial Neural Network* merupakan pendekatan komputasi yang mampu mengikuti cara kerja transfer informasi pada sistem di otak manusia, secara keseluruhan ANN merupakan metode yang mampu diterapkan ketika hubungan antara *input* dan *output* sangat rumit dan *non-linear* [19]. ANN juga dirancang untuk menangani parameter *input* baik yang bersifat kontinu maupun diskrit, agar mampu menangkap interaksi kompleks dalam data [20]. Penyusun ANN terdiri dari tiga lapisan yaitu *input layer* yang berisi sekelompok *neuron input* yang menyediakan masukan untuk ANN, *hidden layer* terdiri dari sekumpulan *neuron* yang tidak terlihat atau tersembunyi, dan *output layer* yang merupakan kumpulan *neuron* yang menjadi keluaran ANN [21]. Pada Gambar 4 merupakan contoh arsitektur milik ANN dengan 4 *input layer*, 6 *hidden layer*, serta 1 *output layer*.



Gambar 4. Arsitektur Artificial Neural Network

Selain layer yang ada pada susunan ANN, ada parameter lain yang berperan penting dalam performa ANN seperti *learning rate* dan fungsi aktivasi. Fungsi aktivasi merupakan elemen yang menentukan rumus matematis yang diterapkan dalam model untuk memproses data *input* agar menghasilkan nilai *output* yang diinginkan, contoh dari fungsi aktivasi adalah *ReLU* yang akan mengaktifkan *neuron* jika hasil dari transformasi *linear* melebihi 0, sehingga menjadi efisien dan mudah dalam perhitungan [19].

2.7 Melatih Dataset Train

Proses pelatihan model dilakukan dengan menggunakan data latih hasil *split dataset* sebelumnya. Model dilatih menggunakan platform *Google Colab* dengan *hardware accelerator T4-GPU* dengan spesifikasi 15 GB GPU RAM, 12.7 GB RAM, dan 112.6 GB Disk. *Hyperparameter* yang digunakan yaitu 50 *epoch* dan 32 *batch size*. *Optimizer* yang akan dipakai yaitu *Adam*, dan fungsi *loss* yang dipakai yaitu *categorical crossentropy* dikarenakan target bersifat multi-kelas. Selama proses pelatihan, model akan menyesuaikan bobot dan bias untuk meminimalkan *error* berdasarkan nilai *loss function* yang dihasilkan.

2.8 Menguji Dataset Test

Setelah proses pelatihan selesai, model terbaik dari hasil *train* akan disimpan untuk kemudian dilakukan proses uji kinerja menggunakan data yang belum pernah dipakai pada proses *training* model.

2.9 Evaluasi Matrik

Tahap selanjutnya adalah evaluasi performa model ANN menggunakan beberapa matrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*.

- Accuracy* berguna untuk mengukur sejauh mana model mengklasifikasikan data dengan benar secara keseluruhan [22]. Pada penelitian ini, *accuracy* digunakan untuk menilai seberapa jauh model ANN dapat mengklasifikasikan dampak (ringan, sedang, berat) gempa bumi ke dalam kategori yang benar secara keseluruhan. Rumus *accuracy*:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} [22]$$

- Precision* berguna mengukur sejauh mana prediksi positif yang akurat [22]. Matrik ini akan digunakan untuk memastikan bahwa prediksi terhadap gempa dengan dampak berat dilakukan secara tepat, hal ini penting agar model tidak menghasilkan *false positive* terhadap gempa yang sebenarnya tidak berdampak berat. Rumus *precision*:

$$Precision = \frac{TP}{(TP+FP)} [22]$$

- c. *Recall* berguna untuk mengukur proporsi dari kasus positif yang sebenarnya dan berhasil diidentifikasi dengan benar [23]. Pada matrik evaluasi ini, digunakan untuk menunjukkan kemampuan model dalam mendeteksi semua kejadian gempa yang benar-benar berdampak berat, semakin tinggi recall maka semakin kecil kemungkinan model melewati kejadian penting. Rumus *recall*:

$$Recall = \frac{TP}{(TP+FN)} [23]$$

- d. *F1-score* merupakan perpaduan dari *precision* dan *recall* yang memberikan ukuran seimbang terhadap performa model [23]. Matrik ini digunakan untuk memberikan gambaran menyeluruh mengenai kinerja model dalam memprediksi dampak gempa secara akurat dan lengkap, terutama ketika data antar kategori tidak seimbang. Rumus *F1-score*:

$$F1 - score = 2 \times \frac{(precision \times recall)}{(precision+recall)} [23]$$

3. Hasil dan Pembahasan

3.1 Hasil Preprocessing

Di tahap ini dilakukan data *cleaning*, yaitu menghapus duplikasi, mengatasi data yang hilang (*missing values*), dan melakukan seleksi fitur. Selanjutnya dilakukan normalisasi terhadap fitur numerik menggunakan metode *Min-Max Scaling*, dengan tujuan agar seluruh fitur berada dalam rentang yang seragam dan meningkatkan efektivitas pelatihan model *Artificial Neural Network* (ANN). Pada Gambar 5 dan 6 menunjukkan dataset setelah dilakukan *preprocessing*.

eventID	datetime	latitude	longitude	magnitude	mag_type	depth	phasecount	azimuth_gap	location	doubleCouple	cid	strikeP1	dipP1	rakeP1	strikeP2	dipP2	rakeP2	azgapR
0	bmg2023abbl 2023-01-02 01:24:33.750	-2534865141	1407386475	541848717	Mw	11	48	1653920571	Near North Coast of West Papua	0.730125	0.269875	61.815582	20.263444	-71.554458	221.443079	70.819571	-96.662820	165.39205
1	bmg2023acqq 2023-01-02 17:40:02.135	5405371666	9698640442	489902015	Mw	14	25	197851796	Northern Sumatra, Indonesia	0.758121	0.241879	154.556586	78.796995	-177.601487	64.084247	87.647217	-11.212579	197.85179
2	bmg2023adpw 2023-01-03 06:55:20.881	-8086607933	1079046707	5012925059	M	24	132	1998681488	Java, Indonesia	0.827734	0.172266	98.502517	71.726934	86.216272	290.493039	18.647592	-191.301777	209.00204
3	bmg2023aejk 2023-01-03 16:32:26.205	-3738483667	1334905396	4898438986	Mw	10	47	408065512	West Papua Region, Indonesia	0.889666	0.110334	92.834829	73.342844	-2.317763	182.699536	87.779547	-163.329963	40.80656
4	bmg2023aghr 2023-01-04 19:16:57.794	-2438380241	1407692871	5212725088	Mw(mB)	10	104	1170746193	Near North Coast of West Papua	0.445954	0.554046	62.832445	55.309421	-126.051638	294.811406	48.335441	-49.628112	174.02781

Gambar 5. Dataset Setelah Mengatasi *Missing Values*

	datetimeFM	latitude	longitude	magnitude	depth	phasecount	azimuth_gap	location	latFM	lonFM	magFM	depthFM	scalarMoment
0	2023-01-01T18:24:30.501282Z	-2534865141	1407386475	541848717	11	48	1653920571	Near North Coast of West Papua	-2.455547	140.782242	5.450264	10	1.421447e+17
1	2023-01-02T10:40:02.13576Z	5405371666	9698640442	489902015	14	25	197851796	Northern Sumatra, Indonesia	5.405372	96.986404	5.113063	10	1.732300e+16
2	2023-01-02T23:55:20.413839Z	-8086607933	1079046707	5012925059	24	132	1998681488	Java, Indonesia	-8.123598	107.902824	4.420491	65	4.528228e+15
3	2023-01-03T09:32:26.187249Z	-3738483667	1334905396	4898438986	10	47	408065512	West Papua Region, Indonesia	-3.736727	133.491547	5.175072	10	2.358623e+16
4	2023-01-04T12:16:58.107803Z	-2438380241	1407692871	5212725088	10	104	1170746193	Near North Coast of West Papua	-2.447020	140.769886	5.134989	29	5.339283e+16

Gambar 6. Dataset Setelah Seleksi Fitur

3.2 Hasil Labeling

Pada tahap ini, dilakukan proses *labeling* untuk menentukan kategori dampak dari gempa bumi berdasarkan kriteria tertentu. Pada penelitian ini, klasifikasi dampak akan dipisah menjadi tiga kategori berbeda seperti kategori ringan, kategori sedang, dan kategori berat. Penentuan label dibuat dengan mempertimbangkan intensitas magnitudo dengan pembagian seperti 0-5.4 untuk ringan, 5.5-6.9 untuk sedang, dan 7.0-8.0 ke atas untuk berat. Kelas ini nantinya digunakan sebagai target *output* dalam proses klasifikasi oleh model ANN. Pada Gambar 7 merupakan hasil setelah dilakukan proses *labeling*.

	datetimeFM	latitude	longitude	magnitude	depth	phasecount	azimuth_gap	location	latFM	lonFM	magFM	depthFM	scalarMoment	magnitude_decimal	label
0	2023-01-01T18:24:30.501282Z	-2534865141	1407386475	541848717	11	48	1653920571	Near North Coast of West Papua	-2.455547	140.782242	5.450264	10	1.421447e+17	5.4	Ringan
1	2023-01-02T10:40:02.13576Z	5405371666	9698640442	489902015	14	25	197851796	Northern Sumatra, Indonesia	5.405372	96.986404	5.113063	10	1.732300e+16	4.8	Ringan
2	2023-01-02T23:55:20.413839Z	-8086607933	1079046707	5012925059	24	132	1998681488	Java, Indonesia	-8.123598	107.902824	4.420491	65	4.528228e+15	5.0	Ringan
3	2023-01-03T09:32:26.187249Z	-3738483667	1334905396	4898438986	10	47	408065512	West Papua Region, Indonesia	-3.736727	133.491547	5.175072	10	2.358623e+16	4.8	Ringan
4	2023-01-04T12:16:58.107803Z	-2438380241	1407692871	5212725088	10	104	1170746193	Near North Coast of West Papua	-2.447020	140.769886	5.134989	29	5.339283e+16	5.2	Ringan

Gambar 7. Dataset Setelah Pelabelan

3.3 Hasil SMOTE

Hasil distribusi kelas dalam data yang telah dilabeli menunjukkan ketidakseimbangan yang signifikan, di mana kelas “berat” memiliki jumlah data yang jauh lebih sedikit dibanding kelas lainnya. Untuk mengatasi masalah ini dan mencegah model menjadi bias terhadap kelas mayoritas, dilakukan teknik *oversampling* terhadap kelas minoritas. Teknik yang digunakan adalah SMOTE (*Synthetic Minority Oversampling Technique*), yang bertujuan untuk menghasilkan data sintesis dari kelas minoritas agar distribusi antar kelas menjadi lebih seimbang. Pada Gambar 8 dan 9 menunjukkan perubahan distribusi kategori antara sebelum dan setelah dilakukan SMOTE.



label	
Ringan	1217
Sedang	94
Berat	7

Gambar 8. Distribusi Kategori Sebelum SMOTE



label	
Ringan	1217
Sedang	500
Berat	50

Gambar 9. Distribusi Kategori Setelah SMOTE

3.4 Hasil Splitting Dataset

Pada tahap ini dilakukan tiga eksperimen pembagian dataset untuk mengetahui jumlah pembagian dataset, yaitu 70% *Training* dan 30% *Testing*, 80% *Training* dan 20% *Testing*, 90% *Training* dan 10% *Testing*. Tujuannya adalah untuk mengetahui konfigurasi pembagian data yang paling optimal bagi performa model. Tiga pembagian dataset ini dipilih karena merupakan pembagian data yang paling umum dalam pengembangan model *machine learning*. Setelah eksperimen dilakukan, rasio pembagian terbaik yang digunakan adalah 70% untuk proses pelatihan dan 30% untuk proses pengujian. Tujuan dari dilakukannya pembagian dataset adalah untuk memastikan jika model ANN mampu secara maksimal belajar dari data train dan diuji performanya terhadap data yang belum pernah dilihat sebelumnya. Pada Tabel 1 merupakan hasil dari tiga pembagian dataset beserta setiap kategorinya.

Tabel 1. Hasil Pembagian Dataset

Pembagian Dataset	Kategori	Accuracy	Precision	Recall	F1-score
70% <i>Training</i> dan 30% <i>Testing</i>	Berat	0.99	0.88	1.00	0.94
	Sedang		0.99	0.99	0.99
	Ringan		1.00	0.99	1.00
80% <i>Training</i> dan 20% <i>Testing</i>	Berat	1.00	0.91	1.00	0.95
	Sedang		1.00	1.00	1.00
	Ringan		1.00	1.00	1.00
90% <i>Training</i> dan 10% <i>Testing</i>	Berat	0.99	0.88	1.00	0.94
	Sedang		0.99	0.99	1.00
	Ringan		1.00	0.99	0.99

Berdasarkan hasil dari eksperimen pembagian dataset, diperoleh pembagian dataset dengan nilai evaluasi terbaik yaitu 80% Train dan 20% Test. Meskipun pembagian 80% Train dan 20% Test menunjukkan hasil evaluasi terbaik dengan akurasi dan F1-score sempurna, namun pada penelitian ini digunakan 70% Train dan 30% Test untuk menghindari *overfitting* dan memastikan model diuji pada jumlah data yang lebih besar.

3.5 Perancangan ANN

Di tahap ini dilakukan perancangan *Artificial Neural Network* dengan tiga lapisan yaitu lapisan *input*, lapisan tersembunyi, dan lapisan *output*. Lapisan *input* memiliki 24 *neuron* yang merepresentasikan 24 bulan dalam dua tahun. Lapisan tersembunyi memiliki dua *layer* yang tersembunyi dengan jumlah *neuron* 32 dan 64 dengan menggunakan fungsi aktivasi *ReLU*. Sedangkan lapisan *output* terdiri dari tiga *neuron* untuk memprediksi label klasifikasi dampak (ringan, sedang, berat), dengan fungsi aktivasi *Softmax*. Pada Gambar 10 merupakan penggalan *code* untuk merancang model ANN.

```
# 11. Model ANN
model = Sequential()
model.add(Dense(64, input_dim=X_train_resampled.shape[1], activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(3, activation='softmax'))
```

Gambar 10. Perancangan Model ANN

3.6 Pelatihan Model ANN

Pada langkah ini dilakukan pelatihan model menggunakan *hyperparameter* 50 *epoch* dan 32 *batch size*. *Optimizer* yang dipakai yaitu *Adam*, dan fungsi *loss* yang dipakai yaitu *categorical_crossentropy* dikarenakan target bersifat multi-kelas. Pada Gambar 11 merupakan penggalan *code* untuk melatih model ANN dan pada Gambar 12 merupakan hasil evaluasi matrik dari pengujian menggunakan data *test*.

```
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
# 12. Latih model (tanpa output epoch)
model.fit(X_train_resampled, y_train_resampled, epochs=50, batch_size=32, validation_split=0.2, verbose=0)
```

Gambar 11. Training Model ANN

Akurasi model: 0.99

Classification Report:

	precision	recall	f1-score	support
Berat	0.88	1.00	0.94	15
Ringan	1.00	0.99	1.00	366
Sedang	0.99	0.99	0.99	150
accuracy			0.99	531
macro avg	0.96	1.00	0.98	531
weighted avg	0.99	0.99	0.99	531

Gambar 12. Hasil Evaluasi Matrik Testing Menggunakan Data Test

Kelas: Berat

TP: 15
TN: 514
FP: 2
FN: 0
Precision: 0.88
Recall: 1.00
F1-score: 0.94

Gambar 13. Hasil Confusion Matrix Kelas Berat

Kelas: Sedang

TP: 147
TN: 378
FP: 3
FN: 3
Precision: 0.98
Recall: 0.98
F1-score: 0.98

Gambar 14. Hasil Confusion Matrix Kelas Sedang

Kelas: Ringan

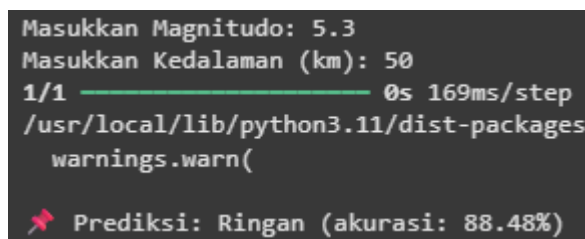
TP: 362
TN: 163
FP: 2
FN: 4
Precision: 0.99
Recall: 0.99
F1-score: 0.99

Gambar 15. Hasil Confusion Matrix Kelas Ringan

Berdasarkan hasil evaluasi model ANN dengan pembagian data 70% *train* dan 30% *test*, diperoleh performa yang sangat baik pada seluruh kelas dampak gempa bumi yang bisa dilihat pada Gambar 13 s.d Gambar 15. Untuk kelas berat, model menghasilkan 15 *True Positive* (TP), 2 *False Positive* (FP), dan 0 *False Negative* (FN) yang menunjukkan bahwa seluruh kejadian berdampak berat berhasil dikenali (*recall* 1.00) meskipun terdapat sedikit kesalahan klasifikasi dari kelas lain (*precision* 0.88). Pada kelas sedang, model menghasilkan 147 TP, 3 FP, dan 3 FN dengan *precision* dan *recall* sebesar 0.98. Sedangkan pada kelas ringan, model berhasil mengklasifikasikan 362 TP dengan hanya 2 FP dan 4 FN, menghasilkan *precision* dan *recall* masing-masing sebesar 0.99. Secara keseluruhan, nilai *F1-score* yang tinggi sebesar 0.94 – 0.99 di ketiga kelas menunjukkan bahwa model memiliki kinerja yang seimbang dan akurat dalam memprediksi tingkat dampak gempa bumi.

3.7 Pengujian Model ANN

Setelah melakukan proses pelatihan dan evaluasi performa, model terbaik disimpan dan digunakan untuk melakukan pengujian prediksi menggunakan fitur Magnitudo dan Kedalaman (*Depth*). Pengujian dilakukan dengan memasukkan jumlah Magnitudo dan *Depth* tertentu, yang mana model harus memprediksi dampak apakah ringan, sedang, atau berat dan dilihat jumlah akurasi prediksinya. Pada Gambar 16 menunjukkan hasil pengujian dengan *input* Magnitudo 5.3 dan *Depth* 50, diperoleh hasil prediksi sebagai dampak ringan dan akurasi sebesar 88.48%.



```
Masukkan Magnitudo: 5.3
Masukkan Kedalaman (km): 50
1/1 _____ 0s 169ms/step
/usr/local/lib/python3.11/dist-packages
warnings.warn(
* Prediksi: Ringan (akurasi: 88.48%)
```

Gambar 16. Hasil Pengujian Model ANN

4. Kesimpulan dan Saran

4.1 Kesimpulan

Berdasarkan dari hasil penelitian yang telah dilakukan, model prediksi dampak gempa bumi di Indonesia berhasil dibangun menggunakan metode *Artificial Neural Network* dan menggunakan *oversampling* SMOTE untuk mengatasi ketidakseimbangan kelas pada *dataset*. Model yang dikembangkan menunjukkan tingkat akurasi sebesar 99% dan nilai *F1-score* yang tinggi dengan kelas berat diperoleh nilai 0.94, kelas sedang diperoleh nilai 0.99, dan kelas ringan diperoleh nilai 1.00. Hasil ini mengindikasikan jika model berhasil melakukan prediksi dengan sangat baik, termasuk pada kelas minoritas seperti kelas berat. Hasil pengujian juga menunjukkan akurasi prediksi yang tetap tinggi sebesar 88.48%.

4.2 Saran

Saran untuk penelitian selanjutnya, dapat menggunakan varian SMOTE lainnya seperti *Borderline-SMOTE* atau *KMeans-SMOTE* untuk mengetahui apakah akurasi dapat ditingkatkan, khususnya pada kelas berat yang jumlahnya terbatas.

Daftar Pustaka

- [1] Y. Syafitri, L. A. Didik, and Bahtiar, “Analisis Pergeseran Lempeng Bumi Yang Meningkatkan Potensi Terjadinya Gempa Bumi Di Pulau Lombok,” Jan. 2020. doi: <https://doi.org/10.20414/konstan.v4i2.43>.
- [2] J. Inayah, A. Fanani, and W. D. Utami, “Klusterisasi Data Kejadian Gempa Bumi di Indonesia Menggunakan Metode K-Medoids,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 2, pp. 271–276, Apr. 2024, doi: 10.26418/justin.v12i2.73594.

- [3] S. Koesuma, V. Fajrin, and B. Sunardi, "Identifikasi Zona Bahaya Gempa Bumi Berdasarkan Percepatan Tanah Maksimum Di Kota Semarang," *Indonesian Journal of Environment and Disaster (IJED)*, vol. 1, no. 2, pp. 65–72, Oct. 2022, doi: <https://doi.org/10.20961/ijed.v1i2.428>.
- [4] A. Fikri, "Badan Geologi Catat Gempa Merusak 2024 Terbanyak dalam 24 Tahun," Bandung, Jan. 2025.
- [5] S. Al Faridzi *et al.*, "Pengolahan Data : Pemahaman Gempa Bumi Di Indonesia Melalui Pendekatan Data Mining," Feb. 2024. doi: <https://doi.org/10.59407/jpki2.v2i1.506>.
- [6] M. Ratih Purwanti, Z. Khalila Salsabila, and F. Liantoni, "Prediksi Gempa Bumi di Yogyakarta Berdasarkan Nilai Magnitudo, Kedalaman, dan Lokasi Gempa Menggunakan Naïve Bayes," *PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 17, no. 1, pp. 122–132, 2024, doi: [10.33322/petir.v17i1.2294](https://doi.org/10.33322/petir.v17i1.2294).
- [7] O. Somantri, S. Purwaningrum, and Riyanto, "Model Support Vector Machine (SVM) Berdasarkan Parameter Windows Untuk Prediksi Kekuatan Gempa Bumi," *Jurnal Teknologi Terapan*, vol. 8, no. 1, pp. 17–24, Mar. 2022, doi: <https://doi.org/10.31884/jtt.v8i1.352>.
- [8] M. A. Permana and M. Faisal, "Uji Performa Prediksi Gempa Bumi di Jawa Timur dengan Artificial Neural Network," *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 11, no. 1, pp. 44–54, May 2023, doi: [10.34312/euler.v11i1.19291](https://doi.org/10.34312/euler.v11i1.19291).
- [9] Suhardjono, G. Wijaya, and A. Hamid, "Prediksi Waktu Kelulusan Mahasiswa Menggunakan SVM Berbasis Pso," *Jurnal Komputer dan Informatika Akademi Bina Sarana Informatika Yogyakarta*, vol. 7, no. 2, pp. 97–101, 2019, doi: <https://doi.org/10.31294/bi.v7i2.6654>.
- [10] M. Agustriya, M. Ula, and Kurniawati, "Analisis Kinerja Algoritma Klasifikasi Naïve Bayes Menggunakan Genetic Algorithm dan Bagging untuk Data Publik Risiko Transaksi Kartu Kredit," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 3, pp. 584–591, Jul. 2024, doi: [10.26418/justin.v12i3.80136](https://doi.org/10.26418/justin.v12i3.80136).
- [11] H. D. Bhakti, "Aplikasi Artificial Neural Network (ANN) untuk Memprediksi Masa Studi Mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik," *Eksplora Informatika*, vol. 9, no. 1, pp. 88–95, Sep. 2019, doi: [10.30864/eksplora.v9i1.234](https://doi.org/10.30864/eksplora.v9i1.234).
- [12] A. K. Vigilant, "Earthquakes in Indonesia," Kaggle.
- [13] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," Sep. 2021. doi: <https://doi.org/10.33884/jif.v9i02.3755>.
- [14] A. S. N. Alarifi, N. S. N. Alarifi, and S. Al-Humidan, "Earthquakes Magnitude Predication Using Artificial Neural Network In Northern Red Sea Area," *J King Saud Univ Sci*, vol. 24, no. 4, pp. 301–313, 2012, doi: [10.1016/j.jksus.2011.05.002](https://doi.org/10.1016/j.jksus.2011.05.002).
- [15] A. Agung, A. Daniswara, I. Kadek, and D. Nuryana, "Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru," *Journal of Informatics and Computer Science*, vol. 05, no. 1, pp. 97–100, 2023, doi: <https://doi.org/10.26740/jinacs.v5n01.p97-100>.
- [16] A. R. Wibowo *et al.*, "Penerapan Metode Naïve Bayes Classifier untuk Klasifikasi Sentimen pada Judul Berita," 2024. doi: <https://doi.org/10.28918/logiclink.v1i1.7684>.
- [17] K. Ujaran, K. Ridwan, E. Heni Hermaliani, M. Ernawati, and C. Author, "Penerapan Metode SMOTE Untuk Mengatasi Imbalanced Data Pada," Jan. 2024. doi: <https://doi.org/10.31294/coscience.v4i1.2990>.
- [18] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," *Jurnal Sains dan Informatika*, vol. 9, no. 1, pp. 19–28, Jun. 2023, doi: [10.34128/jsi.v9i1.622](https://doi.org/10.34128/jsi.v9i1.622).
- [19] R. H. Dananjaya, S. Sutrisno, and S. Fitriady, "Penerapan Artificial Neural Network (ANN) Dalam Memprediksi Kapasitas Dukung Fondasi Tiang," *Matriks Teknik Sipil*, vol. 10, no. 4, pp. 419–426, Dec. 2022, doi: [10.20961/mateksi.v10i4.65034](https://doi.org/10.20961/mateksi.v10i4.65034).

-
- [20] S. Biswas, D. Kumar, M. Nas, M. Softa, E. Akgün, and U. Kumar Bera, “Performance of a Five-Layer ANN Model for Earthquake Magnitude Prediction and Spatial Risk Mapping in Turkey,” *Decision Making Advances Journal homepage: www.dma-journal.org Decision Making Advances*, vol. 3, no. 1, pp. 40–49, 2025, doi: 10.31181/dma31202553.
- [21] A. B. Kurniati, W. A. Sidik, and Jajang, “Model Artificial Neural Networks (ANN) untuk Prediksi COVID-19 di Indonesia,” *JST (Jurnal Sains dan Teknologi)*, vol. 12, no. 3, pp. 833–844, Jan. 2024, doi: 10.23887/jstundiksha.v12i3.53437.
- [22] M. Fadli and R. A. Saputra, “Klasifikasi Dan Evaluasi Performa Model Random Forest Untuk Prediksi Stroke,” *JT: Jurnal Teknik*, vol. 12, no. 2, pp. 72–80, 2023, doi: <http://dx.doi.org/10.31000/jt.v12i2.9099>.
- [23] H. Kaur and D. Kaur Sandhu, “Evaluating the Effectiveness of the Proposed System Using F1 Score, Recall, Accuracy, Precision and Loss Metrics Compared to Prior Techniques,” *International Journal of Communication Networks and Information Security*, vol. 15, no. 4, pp. 368–383, 2023, [Online]. Available: <https://ijcnis.org>