

Penerapan Metode Iteratif Gradient Descent pada Regresi Linear Studi Kasus Data Gempa

Komputasi Statistik

Disusun Oleh : Kelompok Ke - 3

Nama Anggota Kelompok:

- | | |
|--------------------------|-------------|
| 1. Vany Salsabila Putri | (123450022) |
| 2. Kharisma Mustika Sari | (123450034) |
| 3. Aisyah Musfirah | (123450084) |
| 4. Desman Velius Halawa | (123450114) |



PROGRAM STUDI SAINS DATA

FAKULTAS SAINS

INSTITUT TEKNOLOGI SUMATERA

2025

Abstrak

Penelitian ini bertujuan untuk menerapkan pemrograman iteratif melalui algoritma Gradient Descent dalam mengoptimasi parameter model regresi linear untuk memprediksi magnitudo gempa bumi di Southern Sumatera. Data sekunder yang digunakan berasal dari platform Kaggle, mencakup variabel Latitude, Longitude, Depth, dan Magnitude untuk periode 2020–2023. Setelah pra-pemrosesan data dan pembentukan matriks fitur, proses Gradient Descent dilakukan secara iteratif untuk memperbarui parameter model berdasarkan perhitungan prediksi, error, gradien, dan loss pada tiap langkah. Hasil penelitian menunjukkan pola penurunan loss yang tajam pada iterasi awal sebelum melandai mendekati nol, menandakan bahwa proses optimasi berhasil mencapai konvergensi. Trayektori parameter memperlihatkan perubahan besar di awal iterasi yang kemudian stabil, menunjukkan efektivitas mekanisme pembaruan berbasis gradien. Meskipun demikian, evaluasi model menghasilkan MSE sebesar 0.2109, RMSE 0.4592, MAE 0.3620, serta nilai R^2 -0.0364, yang mengindikasikan bahwa regresi linear sederhana belum mampu menangkap hubungan yang kuat antara variabel lokasi dan kedalaman terhadap magnitudo gempa. Secara keseluruhan, penelitian ini menegaskan bahwa algoritma Gradient Descent bekerja efektif sebagai pendekatan iteratif untuk optimasi parameter, meskipun akurasi prediksi tetap bergantung pada relevansi variabel input terhadap fenomena seismik yang kompleks.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan salah satu negara dengan tingkat aktivitas seismik tertinggi di dunia karena berada pada pertemuan tiga lempeng tektonik besar: Indo-Australia, Eurasia, dan Pasifik. Kondisi ini menjadikan wilayah-wilayah di sepanjang jalur subduksi sebagai daerah rawan gempa, termasuk Sumatera bagian selatan (Sumbagsel) seperti Lampung, Bengkulu, Sumatera Selatan, dan Jambi. Kawasan ini terletak pada jalur utama zona subduksi Megathrust Sunda, yang menjadi sumber banyak gempa bumi kuat di Indonesia. Pulau Sumatera sendiri merupakan salah satu wilayah dengan aktivitas seismik paling intens karena berada tepat di atas zona megathrust tersebut. Jalur ini membentang dari pesisir barat Aceh hingga ke selatan Lampung dan dikenal sebagai sumber energi tektonik besar yang kerap memicu gempa signifikan. Pola kegempaan di kawasan ini tidak hanya ditandai oleh frekuensi kejadian yang tinggi, tetapi juga oleh variasi magnitudo dan kedalaman yang cukup besar dari tahun ke tahun.

Data dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) menunjukkan bahwa gempa bumi di wilayah Sumatera Selatan terjadi secara berkala setiap tahun, dengan variasi magnitudo dan kedalaman yang berbeda-beda. Sebagian besar gempa di kawasan ini tergolong gempa dangkal (kedalaman < 60 km), yang cenderung menghasilkan guncangan lebih kuat di permukaan meskipun memiliki magnitudo kecil hingga menengah. Dengan frekuensi gempa yang cukup tinggi, analisis data gempa menjadi penting untuk memberikan gambaran mengenai pola kejadian, hubungan antar variabel, dan potensi risiko di masa mendatang.

Perkembangan teknologi komputasi memberikan peluang besar dalam melakukan prediksi parameter seismik seperti magnitudo dengan pendekatan kuantitatif. Salah satu pendekatan yang dapat digunakan adalah *Gradient Descent*, yaitu algoritma optimasi iteratif yang meminimalkan kesalahan prediksi pada suatu model. Metode ini sering dimanfaatkan dalam regresi linier, machine learning, dan pemodelan statistik karena kemampuannya menemukan parameter optimal melalui proses iterasi berulang. Untuk membangun model prediksi, digunakan algoritma Iterative Gradient Descent, yaitu metode optimasi yang bekerja dengan memperbarui bobot model secara berulang (iteratif) hingga nilai *loss* mengecil. Gradient Descent dipilih karena kemampuannya mengoptimalkan parameter regresi secara efisien dan cocok untuk data berjumlah besar.

1.2 Rumusan Masalah

1. Bagaimana prinsip pemrograman iteratif dapat diterapkan dalam algoritma Gradient Descent untuk mengoptimasi parameter pada model regresi linear?
2. Bagaimana proses pembaruan parameter pada setiap iterasi berpengaruh terhadap penurunan nilai *loss* hingga mencapai konvergensi?

3. Bagaimana performa model regresi linear setelah melalui proses optimasi berbasis metode iteratif Gradient Descent pada studi kasus data gempa Sumatera Selatan?

1.3 Tujuan Penelitian

1. Menerapkan metode pemrograman iteratif melalui algoritma Gradient Descent untuk mengoptimasi model regresi linear.
2. Mengevaluasi performa model dan memahami perilaku konvergensi melalui analisis perubahan parameter beta dan *loss* selama iterasi.

1.4 Manfaat Penelitian

1. Manfaat Praktis

- Memberikan panduan implementasi Gradient Descent bagi mahasiswa atau praktisi yang ingin memahami proses pembaruan bobot secara iteratif.
- Menyediakan contoh aplikasi metode iteratif pada dataset riil (data gempa), sehingga dapat dijadikan acuan untuk penelitian atau pengembangan model sederhana berbasis data.
- Mengilustrasikan bagaimana monitoring *loss* serta parameter dapat membantu pengembang model dalam menentukan stabilitas dan efektivitas proses training.

2. Manfaat Akademis

- Menambah pemahaman mengenai penerapan pemrograman iteratif dalam algoritma optimasi, khususnya Gradient Descent pada regresi linear.
- Memberikan contoh implementasi nyata mengenai bagaimana proses iteratif dapat diobservasi, divisualisasikan, dan dianalisis dalam konteks data numerik.
- Menjadi referensi bagi studi selanjutnya yang membahas optimasi model dan analisis konvergensi metode iteratif.

BAB II

TINJAUAN PUSTAKA

2.1 Gempa Bumi

Gempa bumi adalah peristiwa bergetarnya bumi akibat pelepasan energi dari dalam perut bumi secara tiba-tiba, sehingga menciptakan gelombang seismik, yang ditandai dengan patahnya lapisan batuan pada kerak bumi [3].

Menurut data BMKG (2023–2024), lebih dari seribu kejadian gempa tercatat setiap tahun di wilayah Sumatera dan Sumbagsel. Sebagian besar memiliki magnitudo rendah, namun gempa besar yang bersifat merusak tetap terjadi secara berkala, seperti gempa Aceh 2004, Padang 2009, dan Mentawai 2010. Parameter yang paling sering digunakan untuk menganalisis gempa adalah magnitudo, kedalaman, serta koordinat episenter (lintang dan bujur) [4].

Parameter geospasial seperti *latitude*, *longitude*, dan kedalaman gempa berperan penting dalam memodelkan perilaku dan kekuatan gempa. Beberapa penelitian menunjukkan bahwa data spasial dapat dimanfaatkan untuk estimasi magnitudo menggunakan model statistik dan machine learning [6].

2.2 Regresi Linear dalam Prediksi Magnitudo

Regresi linier memodelkan hubungan antara besarnya gempa bumi yang bergantung- dan satu atau lebih variabel independen dengan menyesuaikan persamaan linier dengan data yang diamati [1]. Pada penelitian prediksi magnitudo gempa, regresi linear sering digunakan sebagai baseline model karena sederhana, stabil, dan mudah diinterpretasikan. Persamaan umum regresi linear berganda adalah:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

dengan $x_1 = \text{latitude}$, $x_2 = \text{longitude}$, dan $x_3 = \text{depth}$

Regresi linear masih menjadi model dasar dalam banyak penelitian prediksi gempa sebelum beralih ke model kompleks seperti *neural network* atau *ensemble learning* [2]. Proses optimasi parameter regresi umumnya dilakukan melalui metode iteratif seperti Gradient Descent.

2.3 Pemrograman Iteratif

Pemrograman iteratif adalah pendekatan penyelesaian masalah menggunakan proses pengulangan (*loop*) yang berjalan hingga hasil mencapai nilai konvergen. Banyak metode komputasi modern menggunakan iterasi.

Metode iteratif lebih efisien digunakan dalam pemrosesan data berskala besar dibanding pendekatan rekursif, terutama dalam perhitungan numerik [7].

2.4 Gradient Descent

Gradient descent merupakan algoritma optimasi iteratif yang digunakan untuk menemukan parameter terbaik pada model dengan meminimalkan fungsi *loss*. Algoritma ini memperbarui parameter secara bertahap ke arah negatif gradien :

$$\beta_j = \beta_j - \alpha \cdot \frac{\partial L}{\partial \beta_j}$$

dengan α sebagai learning rate yang mengontrol besar langkah pembaruan [1].

Gradient descent banyak digunakan dalam metode statistik dan machine learning untuk mempermudah pencarian solusi optimal ketika fungsi loss bersifat kompleks. Penelitian Gencer dan Cizmeci menunjukkan bahwa gradient descent merupakan algoritma inti yang digunakan dalam pelatihan model deep learning untuk estimasi magnitudo secara akurat [5])

2.5 Metrik Evaluasi

Dalam pemodelan regresi (prediksi) magnitudo gempa, diperlukan metrik evaluasi untuk mengukur seberapa baik model dalam mempresentasikan data. Metrik evaluasi yang digunakan :

a. Mean Squared Error (MSE)

MSE adalah ukuran rata-rata selisih kuadrat antara nilai aktual dan nilai prediksi.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad [2]$$

MSE sensitif terhadap outlier sehingga cocok digunakan pada dataset seismik yang memiliki nilai ekstrim.

b. Root Mean Squared Error (RMSE)

Root mean squared error adalah akar kuadrat dari MSE, digunakan untuk melihat seberapa jauh prediksi menyimpang dari nilai aktual dalam satuan yang sama. RMSE cocok digunakan dalam evaluasi prediksi magnitudo karena memberikan interpretasi yang jelas terhadap ketelitian model.

c. Mean Absolute Error (MAE)

Mean absolute error adalah rata-rata kesalahan absolut antara nilai prediksi dan aktual.

$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i| \quad [2]$$

MAE lebih *robust* terhadap outlier dibandingkan MSE. MAE memberikan estimasi error yang lebih stabil pada dataset seismik dengan distribusi tidak normal.

d. Koefisien Determinasi (R^2)

R^2 mengukur seberapa besar variasi target dapat dijelaskan oleh variabel independen. Nilai mendekati 1 menunjukkan model sangat baik. R^2 umum digunakan dalam analisis regresi gempa untuk memahami kontribusi variabel geofisika terhadap magnitudo [2].

BAB III

METODE PENELITIAN

3.1 Jenis Data

Penelitian ini menggunakan data kuantitatif berupa data sekunder yang berasal dari dataset aktivitas gempa bumi. Data tersebut mencakup variabel numerik seperti *Latitude*, *Longitude*, *Depth*, dan *Magnitude* yang masing-masing merepresentasikan posisi geografis, kedalaman pusat gempa, serta kekuatan guncangan. Seluruh variabel bersifat kontinu sehingga dapat diolah menggunakan metode regresi dan algoritma iteratif seperti Gradient Descent. Data sekunder ini tidak diperoleh secara langsung dari lapangan, tetapi dari sumber terbuka (*open data*) yang telah dihimpun dan dipublikasikan sebelumnya.

3.2 Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari platform *Kaggle*, yang menyediakan dataset gempa bumi dengan berbagai atribut terkait aktivitas seismik di Sumatera. Dari dataset tersebut, peneliti melakukan proses seleksi kolom untuk memastikan variabel yang digunakan sesuai dengan tujuan analisis, yaitu *Latitude*, *Longitude*, *Depth/km*, dan *Magnitude*. Data kemudian disimpan dalam *dataframe* dan dipersiapkan untuk tahap pra-pemrosesan serta analisis menggunakan metode pemrograman iteratif.

3.3 Variabel yang Diamati

Isi dataset meliputi informasi-informasi tentang gempa yang pernah terjadi di pulau Sumatera dari tahun 1973-2023. Karena keperluan penelitian, data yang digunakan hanya mencakup tahun 2020-2023 dan lokasi terjadinya gempa adalah Southern Sumatera. Sementara itu, atribut yang digunakan adalah sebagai berikut:

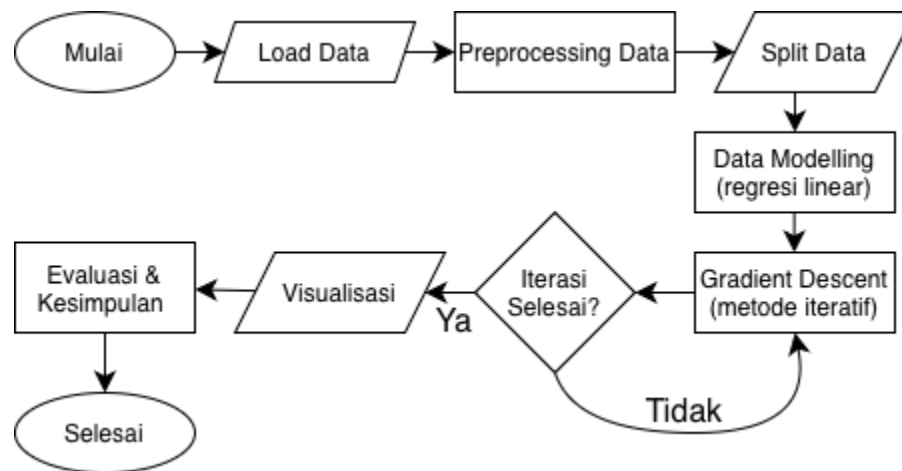
Variabel	Nama Atribut	Tipe Data	Keterangan
Y	Magnitude	Numerik	Magnitudo gempa
X1	Latitude	Numerik	Koordinat lintang lokasi gempa
X2	Longitude	Numerik	Koordinat bujur lokasi gempa
X3	Depth/km	Numerik	Kedalaman pusat gempa dari permukaan (dalam km)

3.4 Diagram Alir Penelitian

Penelitian dimulai dengan mengimpor dataset gempa bumi, kemudian dilanjutkan dengan tahap pra-pemrosesan seperti pemilihan variabel, pembersihan data, serta pembagian data menjadi subset pelatihan dan pengujian. Setelah itu, dilakukan proses normalisasi fitur dan pembentukan matriks input sebagai persiapan untuk pemodelan.

Tahap inti penelitian ditunjukkan pada proses iteratif Gradient Descent, di mana parameter model diinisialisasi terlebih dahulu, kemudian diperbarui secara bertahap melalui perhitungan prediksi, *error*, *gradient*, dan fungsi *loss* pada setiap iterasi. Proses ini berlangsung hingga seluruh iterasi terpenuhi atau model mencapai kondisi konvergensi.

Setelah proses iteratif selesai, dilakukan evaluasi kinerja model menggunakan metrik regresi serta visualisasi seperti kurva *loss* dan tabel perubahan parameter untuk melihat pola optimasi yang terjadi. Diagram alir ini memberikan gambaran menyeluruh mengenai urutan operasional yang dilakukan mulai dari pemrosesan data, proses iteratif, hingga evaluasi hasil akhir.



Gambar 1. Diagram Alir

BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

Penelitian ini menggunakan data sekunder berupa dataset aktivitas gempa bumi yang diperoleh dari platform Kaggle. Dataset berisi variabel numerik Latitude, Longitude, Depth, dan Magnitude, yang merepresentasikan lokasi pusat gempa, kedalaman, dan kekuatan guncangan. Seluruh variabel bersifat kontinu sehingga dapat dianalisis menggunakan regresi dan algoritma iteratif seperti Gradient Descent.

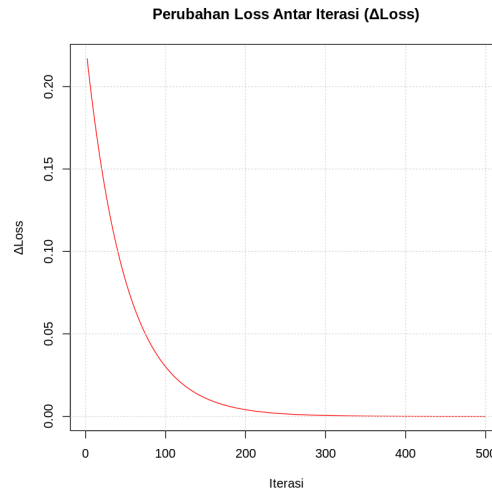
Data asli mencakup aktivitas gempa di Sumatera periode 1973-2023, namun penelitian ini hanya menggunakan subset tahun 2020-2023 dengan fokus wilayah Southern Sumatera. Dataset diekstraksi dan disimpan, kemudian dipilih variabel-variabel yang relevan untuk kebutuhan analisis.

4.2 Hasil Perhitungan

Iterasi	β_0	β_1	β_2	β_3	$\Delta\beta$	Loss (β_j)	ΔLoss
1	0	0	0	0	0.04669081	10.96764	Inf
2	0.04658904	0.001237472	-0.00220683	-0.00175831	0.04622016	10.75073	0.2169044
3	0.09271121	0.00243393	-0.00436085	-0.00348385	0.04575433	10.53818	0.2125537
4	0.13837119	0.003590344	-0.00646319	-0.00517725	0.04529329	10.32989	0.2082911
5	0.18357365	0.004707663	-0.00851500	-0.00683913	0.04483696	10.12577	0.2041147
...
496	4.624216	-0.03334701	-0.08124758	-0.09664292	0.000327324	0.1056494	1.09E-05
497	4.624539	-0.03339245	-0.08127871	-0.09663131	0.000324136	0.1056388	1.07E-05
498	4.624858	-0.03343776	-0.08130989	-0.09661968	0.000320980	0.1056283	1.05E-05
499	4.625174	-0.03348294	-0.08134111	-0.09660802	0.000317857	0.105618	1.03E-05
500	4.625487	-0.03352799	-0.08137237	-0.09659634	0.000314766	0.105608	1.01E-05

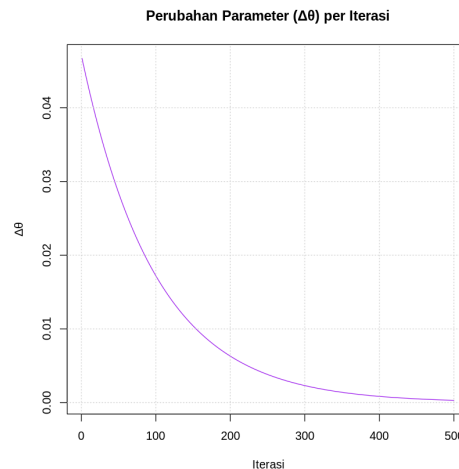
Tabel 1. Hasil Iterasi Gradient Descent pada Model Regresi Linear

Tabel hasil iterasi menunjukkan bagaimana parameter model (β_0, β_1), nilai loss, serta perubahan keduanya ($\Delta\beta$ dan ΔLoss) terus diperbarui di setiap langkah Gradient Descent. Pada iterasi awal, $\Delta\beta$ dan ΔLoss bernilai besar karena algoritma masih melakukan penyesuaian signifikan untuk menurunkan *error*. Namun seiring bertambahnya iterasi, kedua nilai tersebut semakin mengecil dan mendekati nol, menandakan bahwa model bergerak menuju kondisi stabil.



Gambar 1. Perubahan Loss Antar Iterasi (ΔLoss)

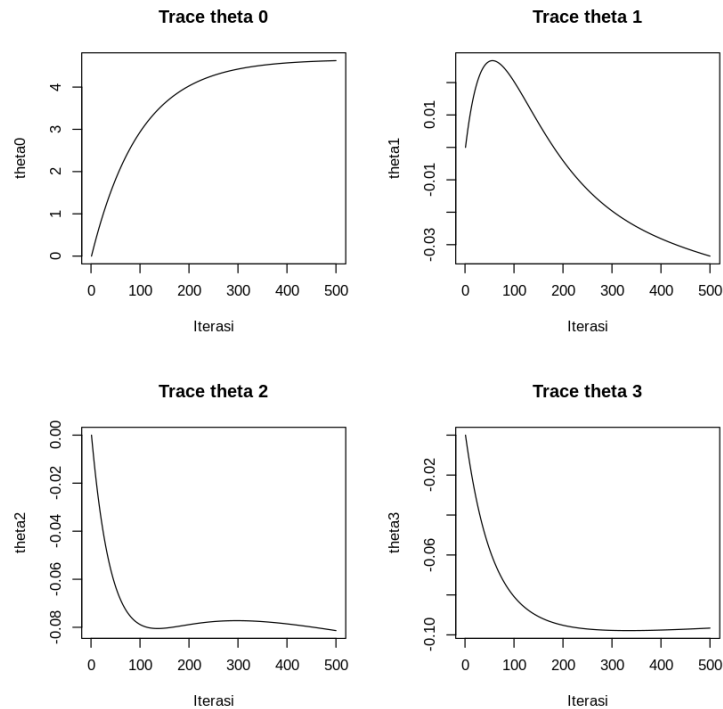
Grafik ΔLoss memperjelas pola tersebut: penurunan loss sangat tajam pada 50–100 iterasi pertama, kemudian melandai hingga hampir datar menjelang iterasi ke-500. Ini mengonfirmasi bahwa proses optimasi berlangsung efektif dan mencapai konvergensi, karena perubahan loss antar iterasi sudah sangat kecil dan parameter tidak lagi berubah secara signifikan.



Gambar 2. Perubahan Parameter β per Iterasi

Grafik $\Delta\beta$ menunjukkan pola konvergensi yang konsisten sepanjang proses Gradient Descent. Pada iterasi awal, nilai $\Delta\beta$ terlihat cukup tinggi, yang menandakan bahwa algoritma melakukan penyesuaian parameter beta untuk memperbaiki error model. Namun seiring bertambahnya iterasi, $\Delta\beta$ terus mengalami penurunan drastis hingga mencapai nilai yang sangat kecil mendekati nol. Pola ini mengindikasikan bahwa setiap

update parameter semakin stabil, karena gradien yang dihitung semakin kecil seiring mendekatnya model ke titik minimum fungsi loss. Dengan kata lain, grafik ini memperlihatkan bukti bahwa proses optimasi berjalan efektif: perubahan parameter besar di awal (fase eksplorasi) bertransisi menjadi perubahan sangat kecil di akhir (fase konvergensi). Hal ini mendukung kesimpulan bahwa model telah mencapai kondisi stabil dan tidak memerlukan pembaruan signifikan lagi.



Gambar 3. Grafik Trayektori Tiap Parameter β

Pada grafik trayektori tiap parameter, setiap garis menunjukkan bagaimana nilai dari satu parameter (β) berubah dari iterasi ke iterasi selama proses gradient descent. Garis-garis tersebut cenderung konvergen ke suatu nilai tertentu, menandakan bahwa algoritma gradient descent telah berhasil mendekati optimal untuk setiap parameter yang meminimalkan fungsi *loss*.

Setelah parameter model mencapai konvergensi, evaluasi performa dilakukan untuk menilai kemampuan model regresi linear dalam memprediksi magnitudo gempa. Nilai MSE sebesar 0.2109, RMSE 0.4592, dan MAE 0.3620 menunjukkan bahwa masih terdapat selisih antara nilai prediksi dan nilai aktual magnitudo. Sementara itu, nilai R^2 sebesar -0.0364 menunjukkan bahwa model belum mampu menjelaskan variasi magnitudo berdasarkan latitude, longitude, dan depth. Hasil ini mengindikasikan bahwa meskipun proses optimasi parameter berjalan dengan baik, hubungan variabel input

terhadap magnitudo gempa tidak cukup kuat untuk direpresentasikan dengan model regresi linear sederhana.

4.3 Interpretasi Hasil

Hasil penelitian menunjukkan bahwa pendekatan pemrograman iteratif melalui algoritma Gradient Descent berjalan efektif dalam mengoptimasi parameter model regresi linear. Pola penurunan loss dan $\Delta\beta$ yang besar di awal, kemudian mengecil hingga mendekati nol, menjadi bukti bahwa proses iteratif bekerja melakukan pembaruan bertahap sampai model mencapai kondisi stabil atau konvergensi.

Meskipun performa prediksi akhir model tidak tinggi, hal ini tidak mengurangi tujuan utama penelitian, yaitu mendemonstrasikan bagaimana mekanisme iteratif bekerja dalam mengoptimasi parameter secara berulang melalui perhitungan gradien. Hasil konvergensi yang jelas menunjukkan bahwa algoritma berhasil menemukan titik minimum lokal dari fungsi loss, bahkan ketika hubungan antar-variabel tidak cukup kuat untuk menghasilkan model prediksi yang akurat.

Dengan demikian, penelitian ini menegaskan bahwa metode iteratif, khususnya Gradient Descent, dapat diterapkan pada data nyata seperti data gempa untuk menunjukkan bagaimana parameter model diperbarui secara sistematis melalui iterasi hingga mencapai kestabilan, sekaligus memberikan gambaran praktis tentang dinamika proses optimasi berbasis gradien.

4.4 Diskusi

Beberapa penelitian terdahulu yang meneliti hubungan parameter subduksi dan magnitudo maksimum menunjukkan bahwa variabel tektonik (mis. slip rate dan karakteristik subduksi) berperan penting dan seringkali perlu dipilih menggunakan teknik *variable selection* sebelum diterapkan pada model regresi. Oleh karena itu, regresi sederhana yang hanya menggunakan variabel lokasi dan kedalaman cenderung berfungsi sebagai *baseline* eksploratori dan dapat melewatkan pengaruh geologi yang lebih substansial. Selain itu, analisis EDA pada dataset gempa global atau regional sebelumnya menunjukkan korelasi yang lemah antara kedalaman dan magnitudo jika variabel tambahan tidak dimasukkan, sehingga interpretasi hubungan ini harus dilakukan dengan hati-hati. Di sisi lain, literatur mutakhir tentang aplikasi *machine learning* pada prediksi magnitudo melaporkan peningkatan performa ketika metode ML (mis. *ensemble* atau jaringan saraf) diaplikasikan pada dataset yang kaya atribut; namun demikian metode tersebut memerlukan data yang lebih besar, pemrosesan fitur yang matang, serta validasi temporal/spasial yang ketat. Penelitian ini menempatkan model regresi iteratif (Gradient Descent) sebagai baseline empiris untuk periode 2020-2023 di Southern Sumatera, dan merekomendasikan penelitian lanjutan yang menggabungkan parameter tektonik dan pendekatan ML untuk meningkatkan pemahaman dan akurasi prediksi.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini bertujuan untuk menerapkan metode pemrograman iteratif melalui algoritma *Gradient Descent* dalam proses optimasi parameter pada model regresi linear menggunakan data gempa bumi wilayah Southern Sumatera. Berdasarkan hasil analisis dan proses iteratif yang dilakukan hingga 500 iterasi, dapat disimpulkan bahwa algoritma *Gradient Descent* berhasil menunjukkan pola optimasi yang sesuai teori, yaitu penurunan nilai *loss* yang konsisten dan $\Delta\beta$ yang semakin kecil seiring meningkatnya jumlah iterasi. Hal ini membuktikan bahwa mekanisme iteratif bekerja dengan stabil dan model mencapai kondisi konvergensi.

Tabel iterasi dan grafik ΔLoss menunjukkan bahwa penurunan kesalahan sangat signifikan pada awal proses, kemudian melandai hingga mendekati nol di akhir iterasi. Pola ini sejalan dengan karakteristik metode *Gradient Descent*, di mana pembaruan parameter dilakukan secara agresif di awal lalu stabil ketika mendekati titik minimum dari fungsi *loss*. Visualisasi trayektori parameter juga menunjukkan bahwa nilai β_0 , β_1 , β_2 dan β_3 bergerak menuju nilai yang konvergen.

Meskipun proses optimasi berjalan efektif, performa prediktif model regresi linear tidak menunjukkan hasil yang kuat dalam menjelaskan variasi magnitudo gempa. Nilai MSE sebesar 0.2109, RMSE sebesar 0.4592, MAE sebesar 0.3620, serta nilai R^2 yang negatif (-0.0364) mengindikasikan bahwa hubungan antara variabel *latitude*, *longitude*, dan *depth* terhadap magnitudo tidak cukup kuat untuk direpresentasikan dengan model regresi linear sederhana. Dengan demikian, penelitian ini berhasil mendemonstrasikan penerapan pemrograman iteratif untuk optimasi model, namun model regresi linear belum mampu menangkap pola hubungan variabel secara akurat pada dataset gempa ini.

5.2 Saran

Berdasarkan hasil yang diperoleh, terdapat beberapa saran yang dapat dilakukan untuk penelitian selanjutnya:

1. Gunakan model yang lebih kompleks

Karena nilai R^2 negatif menandakan regresi linear kurang sesuai, disarankan untuk mengeksplorasi model yang mampu menangkap hubungan non-linier seperti *Random Forest*, *Gradient Boosting*, *Neural Network*, atau *Support Vector Regression*.

2. Lakukan eksplorasi fitur tambahan

Variabel yang digunakan (*latitude*, *longitude*, *depth*) mungkin belum mencukupi. Penelitian selanjutnya dapat menambahkan variabel seperti jenis patahan, jarak ke zona subduksi, atau parameter geofisika lainnya yang lebih relevan dengan magnitudo.

3. Lakukan analisis normalisasi dan parameter tuning lebih lanjut
Hasil optimasi mungkin dipengaruhi oleh pemilihan *learning rate* tunggal. Eksperimen dengan berbagai nilai *alpha*, *batch size*, atau metode optimasi lain seperti *momentum*, *AdaGrad*, atau *Adam* dapat meningkatkan hasil.
4. Gunakan dataset yang lebih besar dan lebih bersih
Data gempa yang digunakan terbatas pada beberapa tahun terakhir dan wilayah tertentu. Penggunaan dataset dengan rentang waktu lebih panjang atau praproses data yang lebih ketat dapat meningkatkan akurasi model.
5. Lakukan evaluasi visual tambahan
Seperti *scatter plot* prediksi vs aktual, *residual plot*, atau *heatmap korelasi* untuk memperdalam interpretasi model dan memastikan apakah masalah utama terletak pada data, model, atau parameter.
Dengan memperhatikan saran-saran tersebut, penelitian lanjutan dapat menghasilkan model prediksi yang lebih akurat serta memberikan pemahaman yang lebih baik mengenai hubungan antar variabel dalam data kegempaan.

BAB VI

DAFTAR PUSTAKA


- [1] S. M. El-Salhi, Bashar Igried, and S. Awwad, “Advanced Machine Learning Methodology for Earthquake Magnitude Forecasting Using Comprehensive Seismic Data,” *Journal of Soft Computing in Civil Engineering*, vol. 10, no. 1, pp. e223293-, 2026, doi: <https://doi.org/10.22115/scce.2025.1955>.
- [2] Krishna Bhargavi Y *et al.*, “Earthquake detection and early warning prediction using folium and Geopandas,” *Cogent engineering*, vol. 11, no. 1, May 2024, doi: <https://doi.org/10.1080/23311916.2024.2345301>.
- [3] I. Kaftan, “Machine Learning Applications for Earthquake Magnitude Prediction in Western Turkey,” *Applied Sciences*, vol. 15, no. 20, pp. 10909–10909, Oct. 2025, doi: <https://doi.org/10.3390/app152010909>.
- [4] BMKG, “Data Gempabumi Terbuka BMKG,” *Bmkg.go.id*, 2025. <https://data.bmkg.go.id/gempabumi/> (accessed Jun. 08, 2025).
- [5] K. GENCER and I. H. CIZMECI, “Enhanced Earthquake Magnitude Prediction Using Hybrid Machine Learning and Deep Learning Models,” <https://dergipark.org.tr/tr/download/article-file/4715683>, 2025. <https://dergipark.org.tr/tr/download/article-file/4715683>
- [6] H. Utama, A. Masruro, T. Indriyatmoko, and Sudarmanto, “Integrasi Augmentasi Data dan Machine Learning dalam Prediksi Magnitudo Gempa Bumi,” *Jurnal Multidisipin*, vol. 2, no. 3, pp. 97–108, 2025.
- [7] M. Endres, W. Weimer, and A. Kamil, “An Analysis of Iterative and Recursive Problem Performance,” *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, vol. -, Mar. 2021, doi: <https://doi.org/10.1145/3408877.3432391>.

LAMPIRAN

Lampiran 1. Sumber Data

- Dataset Gempa Kaggle:
<https://www.kaggle.com/datasets/fikridwialpian/earthquake-in-sumatera>
- Dataset yang telah disortir:  data_gempa

Lampiran 2. Kode Pengerjaan

-  (2) Komstat Tubes.ipynb