

**PERBANDINGAN ALGORITMA RANDOM FOREST DAN XGBOOST
REGRESSOR DALAM KASUS REGRESI UNTUK MEMPREDIKSI
KEKUATAN GEMPA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat

Sarjana

Program Studi informatika



disusun oleh

ANDRIAN PRADANA

21.11.4082

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

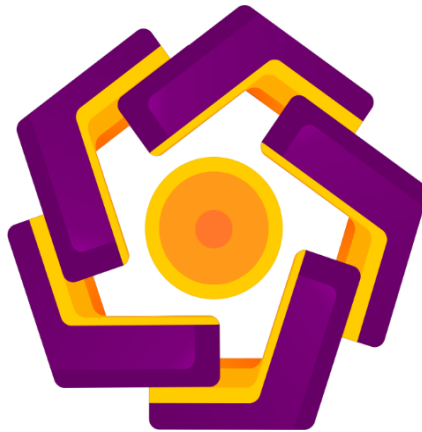
**PERBANDINGAN ALGORITMA RANDOM FOREST DAN XGBOOST
REGRESSOR DALAM KASUS REGRESI UNTUK MEMPREDIKSI
KEKUATAN GEMPA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat

Sarjana

Program Studi informatika



disusun oleh

ANDRIAN PRADANA

21.11.4082

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**PERBANDINGAN ALGORITMA RANDOM FOREST DAN
XGBOOST REGRESSOR DALAM KASUS REGRESI UNTUK
MEMPREDIKSI KEKUATAN GEMPA SKRIPSI**

yang disusun dan diajukan oleh

Andrian Pradana

21.11.4082

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 9 juni 2025

Dosen Pembimbing,



Hastari Utama, S.Kom., M.Cs.
NIK. 190302230

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN ALGORITMA RANDOM FOREST DAN XGBOOST REGRESSOR DALAM KASUS REGRESI UNTUK MEMPREDIKSI KEKUATAN GEMPA SKRIPSI

yang disusun dan diajukan oleh

Andrian Pradana

21.11.4082

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ainul Yaqin, S.Kom., M.Kom.
NIK. 190302255

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

Hastari Utama, S.Kom., M.Cs.
NIK. 190302230

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 juli 2025

DEKAN FAKULTAS KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : ANDRIAN PRADANA
NIM : 21.11.4082

Menyatakan bahwa Skripsi dengan judul berikut:

PERBANDINGAN ALGORITMA RANDOM FOREST DAN XGBOOST REGRESSOR DALAM KASUS REGRESI UNTUK MEMPREDIKSI KEKUATAN GEMPA

Dosen Pembimbing : Hastari Utama, S.Kom., M.Cs.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

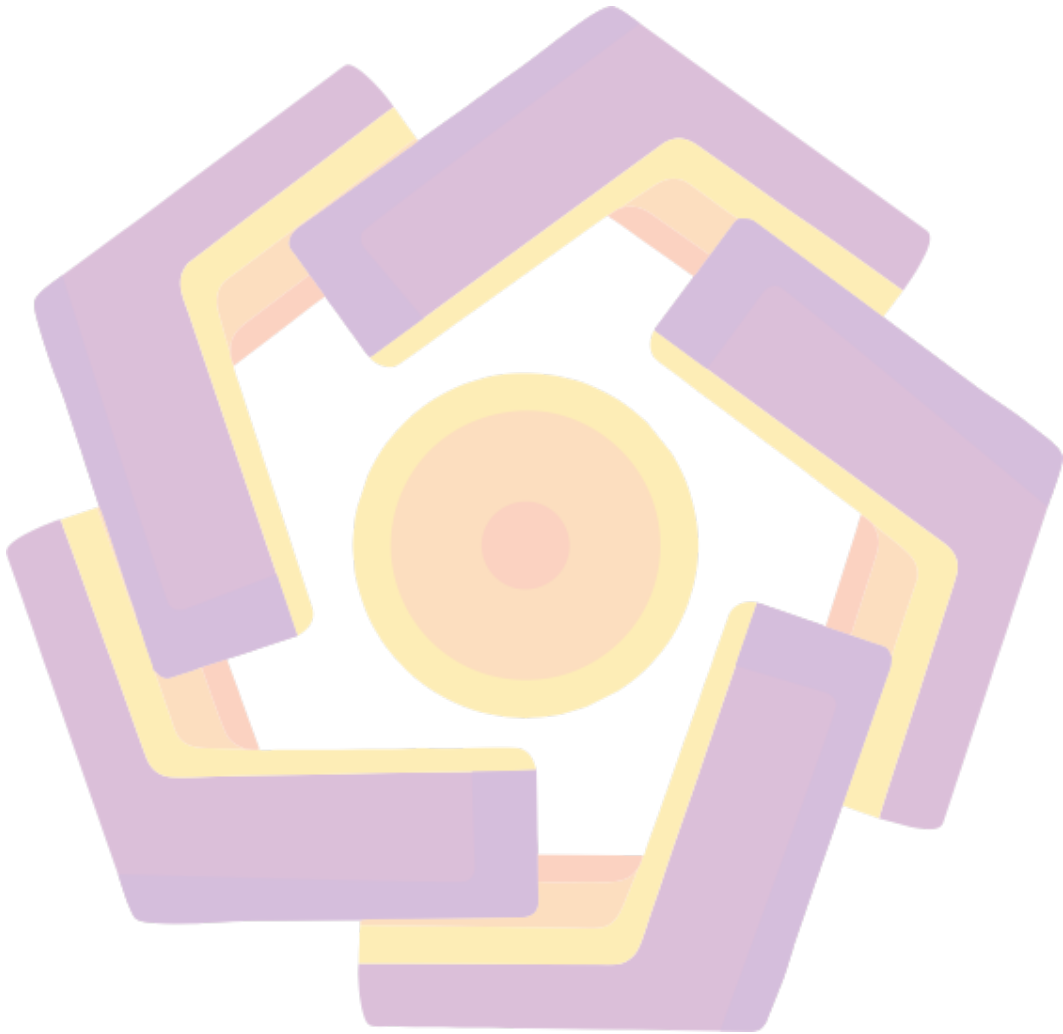
Yogyakarta, 21/07/2025



ANDRIAN PRADANA

HALAMAN PERSEMBAHAN

(Bila ada) Halaman ini berisi kepada siapa skripsi dipersembahkan. Ditulis dengan singkat, resmi, sederhana, tidak terlalu banyak, serta tidak menjurus ke penulisan informal sehingga mengurangi sifat resmi laporan ilmiah.



KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul “Perbandingan algoritma Random Forest dan Xgboost regressor dalam kasus regresi untuk memprediksi kekuatan gempa”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi [isi nama program studi], Fakultas [isi nama fakultas], [isi nama universitas].

Penyusunan skripsi ini tidak lepas dari dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih dan penghargaan yang setinggi-tingginya kepada:

- Yth. Bapak Hastari Utama, S.Kom., M.Cs. selaku Dosen Pembimbing, yang telah dengan sabar membimbing, memberikan arahan, serta meluangkan waktu dalam proses penyusunan skripsi ini.
- Yth. Bapak/Ibu Dosen Penguji dan seluruh dosen di lingkungan [nama fakultas/program studi] yang telah memberikan ilmu dan bimbingannya selama masa studi.
- Kepada kedua orang tua tercinta, Samsul Hadi dan Winnarsih, atas kasih sayang, doa, serta dukungan moril dan materiil yang tak pernah berhenti selama proses perkuliahan dan penyusunan skripsi ini.
- Kepada teman-teman dan semua pihak yang telah membantu dan memberikan semangat selama proses pengerjaan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat dan menjadi referensi bagi pembaca yang membutuhkan.

Yogyakarta, 19 Juli 2025

Penulis

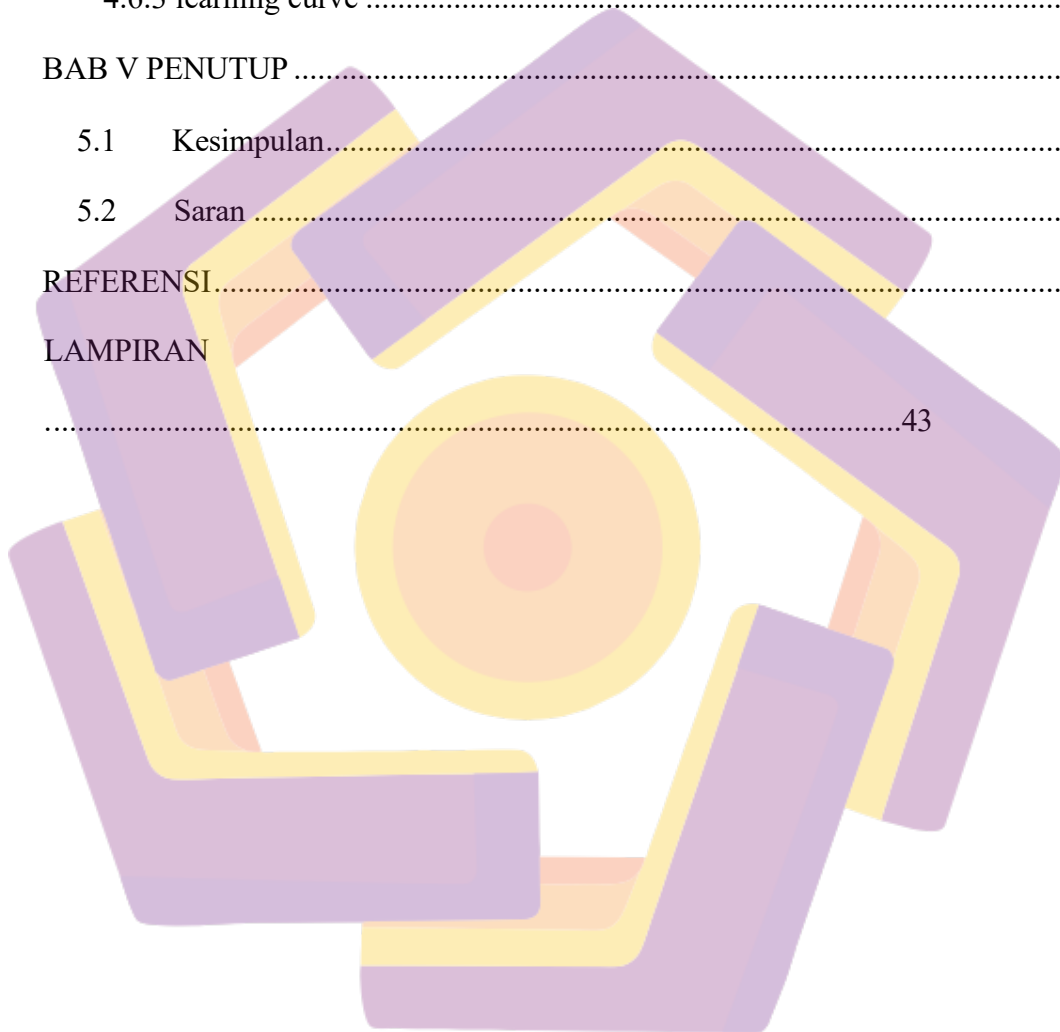
DAFTAR ISI

Contents

HALAMAN PENGESAHAN	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSEMBAHAN.....	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH	xii
INTISARI.....	xiii
ABSTRACT.....	xiv
 BAB I PENDAHULUAN.....	 1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	4
 BAB II TINJAUAN PUSTAKA	 7
2.1 Studi Literatur.....	7

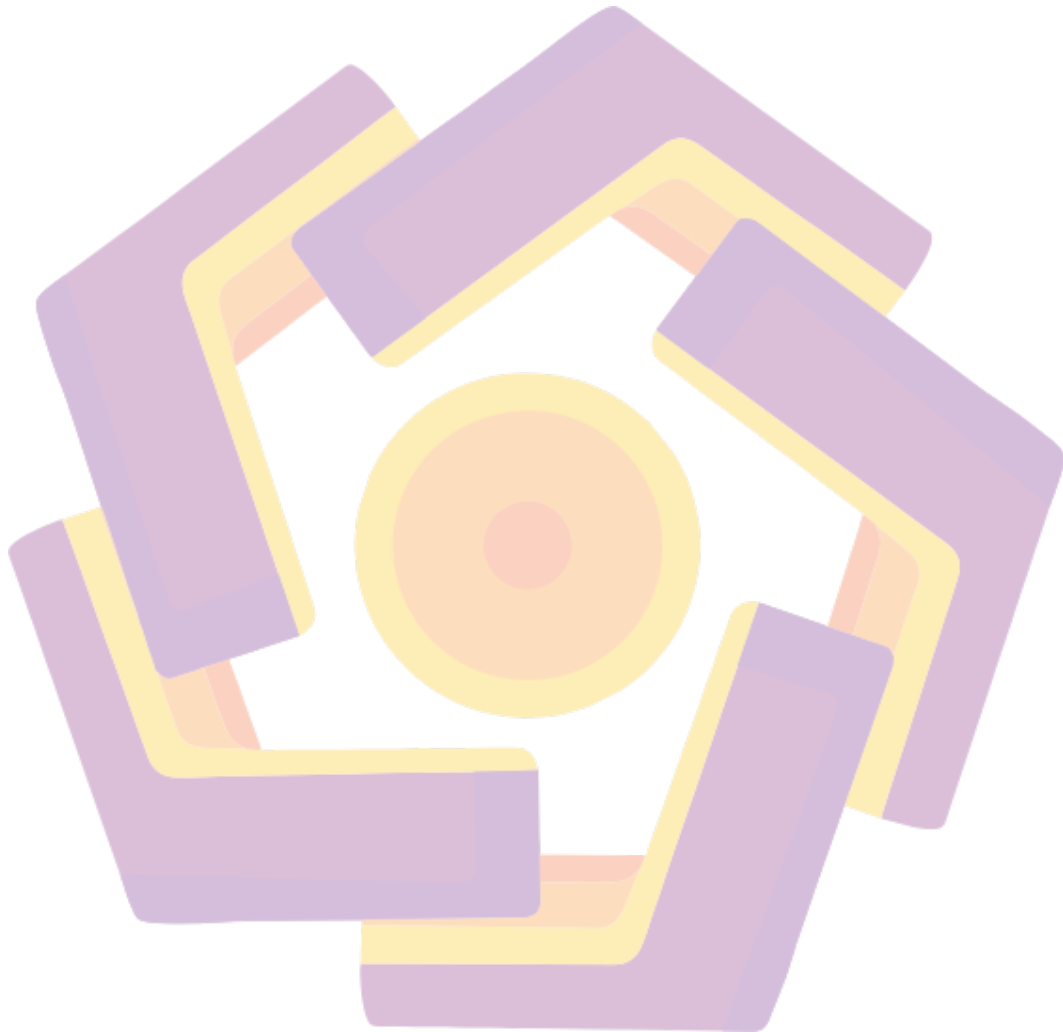
2.2	Dasar Teori	13
2.2.1	Prediksi.....	13
2.2.2	Google colab.....	13
2.2.3	Python.....	14
2.2.4	Gempa	14
2.2.5	Kaggle.....	15
2.2.6	Mechine learning.....	16
2.2.7	Supervised Learning.....	16
2.2.8	Ensemble Learning.....	16
2.2.9	Algoritma Random Forest.....	17
2.2.10	Algoritma Xgtoost.....	19
2.2.11	Cara kerja algoritma XGBoost?.....	19
2.2.12	Regresi linear.....	21
2.2.13	Data seismik.....	21
BAB III	METODE PENELITIAN	22
3.1	Objek Penelitian.....	22
3.2	Alur Penelitian.....	22
3.2	Alur Penelitian.....	23
3.3	Alat dan Bahan.....	26
3.3.1	Data Penelitian.....	26
3.3.2	Alat/instrumen.....	27
BAB IV	HASIL DAN PEMBAHASAN	28
4.1	Load data.....	28
4.2	EDA	29
4.3	Korelasi.....	30
4.4	Spliting data.....	31
4.5	Bangun model.....	32

4.5.1	Random forest.....	32
4.5.2	Xgboost.....	34
4.6	Evaluasi	35
4.6.1	Hasil dari prediksi algoritma.....	35
4.6.2	Setelah dilakukan Hyperparameter Tuning.....	35
4.6.3	learning curve	36
BAB V PENUTUP		38
5.1	Kesimpulan.....	38
5.2	Saran	38
REFERENSI.....		40
LAMPIRAN		43



DAFTAR TABEL

Tabel 2. 1 Keaslian Penulisan.....	10
Tabel 4. 1 Hasil Prediksi Algoritma.....	35
Tabel 4. 2 Hasil Prediksi Setelah hyperparameter tuning Algoritma.....	35



DAFTAR GAMBAR

Gambar 2. 1	13
Gambar 2. 2	14
Gambar 2. 3	15
Gambar 2. 4	15
Gambar 3. 1	23
Gambar 3. 2	
Gambar 4. 1	28
Gambar 4. 2	29
Gambar 4. 3	30
Gambar 4. 4	31
Gambar 4. 5	32
Gambar 4. 6	34
Gambar 4. 7	36
Gambar 4. 8	37

DAFTAR LAMPIRAN

DAFTAR LAMBANG DAN SINGKATAN

Lambang/Singkatan	Keterangan
MAE	Mean Absolute Error (Rata-rata Kesalahan Absolut)
MSE	Mean Squared Error (Rata-rata Kuadrat Kesalahan)
RMSE	Root Mean Squared Error (Akar Kuadrat Rata-rata Kesalahan)
R^2	Koefisien Determinasi
EDA	Exploratory Data Analysis (Analisis Data Eksploratori)
ML	Machine Learning
RF	Random Forest
GB	Gradient Boosting
XGB	Extreme Gradient Boosting (XGBoost)
CV	Cross Validation (Validasi Silang)
API	Application Programming Interface

DAFTAR ISTILAH

Istilah	Definisi
Algoritma masalah.	Langkah-langkah sistematis dalam menyelesaikan suatu
Regresi	Pendekatan statistik untuk memprediksi nilai numerik berdasarkan variabel input.
Random Forest	Algoritma ensemble yang membangun banyak pohon keputusan untuk prediksi.
Gradient Boosting	gabungan model-model lemah secara bertahap.
Prediksi	Perkiraan nilai keluaran berdasarkan data input.
Evaluasi Model	Proses menilai performa model menggunakan metrik seperti MAE, MSE, dan R^2 .
Overfitting	Ketika model terlalu menyesuaikan diri dengan data
Hyperparameter Tuning	Proses mencari kombinasi parameter terbaik untuk meningkatkan performa model.
Validasi Silang (Cross Validation)	Teknik evaluasi model dengan membagi data menjadi beberapa subset untuk pelatihan dan pengujian.

INTISARI

Indonesia merupakan salah satu negara dengan tingkat aktivitas seismik tertinggi di dunia karena berada di pertemuan tiga lempeng tektonik utama. Akibatnya, gempa bumi kerap terjadi dan menimbulkan dampak serius terhadap masyarakat, infrastruktur, dan lingkungan. Tantangan utama dalam penanggulangan bencana gempa bumi adalah kesulitan dalam memprediksi kekuatan gempa secara akurat, mengingat kompleksitas data seismik yang tinggi. Oleh karena itu, dibutuhkan metode yang mampu mengolah data dengan baik dan menemukan pola tersembunyi untuk membantu dalam proses prediksi.

Penelitian ini bertujuan untuk mengimplementasikan dan membandingkan dua algoritma machine learning, yaitu Random Forest dan Gradient Boosting, dalam melakukan prediksi kekuatan gempa bumi. Langkah-langkah yang dilakukan meliputi pengumpulan data gempa dari sumber terbuka, pembersihan data (data cleaning), analisis data secara eksploratori, normalisasi, pemisahan data latih dan uji, pelatihan model menggunakan kedua algoritma, serta evaluasi kinerja model menggunakan metrik Mean Squared Error (MSE), Mean Absolute Error (MAE), dan koefisien determinasi (R^2).

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem prediksi kekuatan gempa berbasis data historis. Hasil dari penelitian ini dapat dimanfaatkan oleh para peneliti, institusi kebencanaan, serta pengembang sistem mitigasi bencana untuk mendukung pengambilan keputusan yang lebih tepat. Penelitian lebih lanjut disarankan untuk mempertimbangkan variabel tambahan seperti struktur geologi dan jenis patahan guna meningkatkan akurasi model.

Kata kunci: gempa bumi, machine learning, random forest, gradient boosting, prediksi.

ABSTRACT

Indonesia is one of the countries with the highest seismic activity in the world due to its location at the convergence of three major tectonic plates. As a result, earthquakes occur frequently and can have serious impacts on society, infrastructure, and the environment. One of the main challenges in disaster mitigation is the difficulty in accurately predicting earthquake magnitude due to the complexity of seismic data. Therefore, there is a need for methods capable of efficiently processing large volumes of data and detecting hidden patterns to support the prediction process.

This study aims to implement and compare two machine learning algorithms, namely Random Forest and Gradient Boosting, in predicting earthquake magnitude. The research steps include collecting earthquake data from open-source platforms, performing data cleaning, conducting exploratory data analysis, normalizing the data, splitting it into training and testing sets, training the models using both algorithms, and evaluating the models using performance metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Rsquared (R^2).

This research is expected to contribute to the development of earthquake magnitude prediction systems based on historical data. The results may be utilized by researchers, disaster management agencies, and developers of disaster mitigation systems to support more accurate and data-driven decision-making. Future studies are recommended to include additional variables such as geological structure and fault types to improve prediction accuracy.

Keywords: earthquake, machine learning, random forest, gradient boosting, prediction.