

Prediksi Magnitudo Gempa Menggunakan Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron Berdasarkan Data Kedalaman dan Geolokasi

Ika Maulita¹, Arif Mu'amar Wahid^{*2}

¹Jurusan Fisika, Fakultas MIPA, Universitas Jenderal Soedirman, Indonesia

²Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, Indonesia

Email: ikamaulita@gmail.com, arif@amikompurwokerto.ac.id

Abstrak

Penelitian ini bertujuan membandingkan kinerja lima algoritma pembelajaran mesin, yaitu Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron dalam memprediksi magnitudo gempa berdasarkan data kedalaman dan geolokasi. Masalah yang diangkat adalah pentingnya prediksi magnitudo gempa yang lebih akurat untuk meningkatkan efektivitas mitigasi risiko bencana, terutama di daerah rawan gempa. Data yang digunakan mencakup informasi kedalaman, lintang, dan bujur dari peristiwa gempa selama periode tertentu. Metode penelitian melibatkan pembagian data pelatihan dan pengujian, serta evaluasi kinerja model menggunakan metrik Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R^2 . Hasil penelitian menunjukkan bahwa LightGBM memberikan performa terbaik dengan nilai MAE sebesar 0,4688, RMSE sebesar 0,6284, dan R^2 sebesar 0,2458. Random Forest mengikuti dengan nilai MAE sebesar 0,4750, RMSE sebesar 0,6312, dan R^2 sebesar 0,2391. XGBoost menunjukkan performa yang kompetitif dengan MAE sebesar 0,4932, RMSE sebesar 0,6471, dan R^2 sebesar 0,2003. Sebaliknya, Support Vector Regression mencatatkan nilai MAE sebesar 0,5136, RMSE sebesar 0,6987, dan R^2 sebesar 0,0677, sementara Multi-Layer Perceptron memberikan kinerja terendah dengan MAE sebesar 0,5190, RMSE sebesar 0,7152, dan R^2 sebesar 0,0231. Dampak penelitian ini sangat penting bagi pengembangan sistem peringatan dini gempa dan peningkatan akurasi prediksi magnitudo gempa. Penelitian ini menegaskan bahwa pemilihan model yang tepat dapat berkontribusi pada mitigasi risiko bencana, dengan memberikan informasi yang lebih akurat mengenai kekuatan gempa yang dapat terjadi. Temuan ini juga menunjukkan bahwa algoritma pembelajaran mesin, terutama LightGBM dan Random Forest, dapat menjadi alat yang efektif dalam analisis seismologi dan aplikasi prediksi gempa.

Kata kunci: *lightgbm, multi-layer perceptron, prediksi gempa, random forest, support vector regression, xgboost*

Predicting Earthquake Magnitude Using Random Forest, Support Vector Regression, XGBoost, LightGBM, and Multi-Layer Perceptron Based on Depth and Geolocation Data

Abstract

This study aims to compare the performance of five machine learning algorithms, namely Random Forest, Support Vector Regression, XGBoost, LightGBM, and Multi-Layer Perceptron, in predicting earthquake magnitudes based on depth and geolocation data. The problem addressed is the importance of more accurate earthquake magnitude predictions to improve the effectiveness of disaster risk mitigation, particularly in earthquake-prone areas. The data used includes information on depth, latitude, and longitude from earthquake events over a specified period. The research method involves dividing the data into training and testing sets, followed by model performance evaluation using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 metrics. The results show that LightGBM provides the best performance with an MAE of 0.4688, RMSE of 0.6284, and R^2 of 0.2458. Random Forest follows with an MAE of 0.4750, RMSE of 0.6312, and R^2 of 0.2391. XGBoost shows competitive performance with an MAE of 0.4932, RMSE of 0.6471, and R^2 of 0.2003. In contrast, Support Vector Regression recorded an MAE of 0.5136, RMSE of 0.6987, and R^2 of 0.0677, while Multi-Layer Perceptron achieved the lowest performance with an MAE of 0.5190, RMSE of 0.7152, and R^2 of 0.0231. The impact of this study is crucial for the development of earthquake early warning systems and improving earthquake magnitude prediction accuracy. This research confirms that choosing the right model can contribute to disaster risk mitigation by providing more accurate information about potential earthquake

magnitudes. The findings also show that machine learning algorithms, particularly LightGBM and Random Forest, can be effective tools in seismology analysis and earthquake prediction applications.

Keywords: *earthquake prediction, lightgbm, multi-layer perceptron, random forest, support vector regression, xgboost*

1. PENDAHULUAN

Gempa bumi merupakan salah satu bencana alam yang memiliki dampak signifikan terhadap masyarakat, infrastruktur, dan perekonomian. Antara tahun 2000 hingga 2019, gempa bumi telah menyebabkan lebih dari 721.000 kematian di seluruh dunia, yang menunjukkan potensi bencana ini untuk menghancurkan kehidupan dan ekonomi secara masif [1]. Dampak ekonomi dari gempa bumi juga sangat besar; misalnya, gempa Sichuan tahun 2008 menyebabkan kerugian langsung sebesar sekitar \$36,6 miliar, yang memperlihatkan betapa rentannya sistem infrastruktur terhadap bencana seismik. Selain kehancuran fisik yang ditimbulkan secara langsung, dampak sosial dari gempa bumi juga tidak kalah penting. Manajemen risiko bencana yang efektif sangat diperlukan untuk meminimalkan jumlah korban jiwa dan kerugian ekonomi. Hal ini mencakup metode penilaian risiko yang dapat mendukung perencanaan kota dan meningkatkan ketahanan infrastruktur. Sebagai contoh, model cerdas hibrida yang menggunakan optimasi swarm partikel telah dikembangkan untuk menilai risiko seismik perkotaan, membantu pemerintah dalam mengelola bencana [2]. Lebih lanjut, persepsi publik dan kesiapan masyarakat berperan penting dalam mengurangi dampak gempa bumi. Penelitian menunjukkan bahwa meningkatkan persepsi risiko melalui edukasi dapat secara signifikan memengaruhi niat individu untuk mempersiapkan diri menghadapi gempa bumi [3].

Chile menjadi fokus penting dalam kajian seismik karena posisinya di Cincin Api Pasifik, kawasan dengan aktivitas seismik tinggi yang sering memicu gempa bumi dan letusan gunung berapi akibat pergerakan lempeng tektonik. Beberapa gempa bumi paling kuat dalam sejarah terjadi di Chile, termasuk gempa Valdivia tahun 1960 yang tercatat sebagai gempa terkuat dengan magnitudo 9,5 [4]. Gempa-gempa ini tidak hanya menyebabkan kerusakan langsung tetapi juga meninggalkan dampak jangka panjang terhadap komunitas dan infrastrukturnya [5]. Kondisi geologis Chile membuatnya rentan terhadap bahaya seismik. Negara ini berada di pertemuan lempeng Nazca dan Amerika Selatan, di mana lempeng Nazca mengalami subduksi di bawah lempeng Amerika Selatan. Proses subduksi ini menyebabkan aktivitas seismik yang signifikan, termasuk gempa bumi megathrust dan tsunami terkait [6]. Sebagai contoh, gempa bumi di Constitución pada tahun 2010 menyebabkan kerusakan yang luas dan banyak korban jiwa, yang menunjukkan urgensi pengembangan strategi kesiapsiagaan dan tanggap bencana yang efektif [7]. Frekuensi historis kejadian seismik di Chile menunjukkan perlunya penelitian berkelanjutan mengenai prediksi gempa bumi dan mitigasi risiko, serta pengembangan infrastruktur yang tahan gempa.

Prediksi dan pemodelan gempa bumi telah mengalami perkembangan yang signifikan seiring dengan kemajuan dalam pengumpulan data dan metodologi statistik. Salah satu aspek penting dari perkembangan ini adalah integrasi berbagai sumber data seismik, termasuk pengukuran GPS, yang terbukti meningkatkan akurasi prediksi gempa bumi. Wang et al. menunjukkan bahwa penggunaan data GPS dalam model probabilistik, terutama yang menyerupai proses Hawkes yang terstimulasi sendiri, dapat memperbaiki prediksi gempa jangka pendek dengan memberikan pemahaman yang lebih mendalam tentang aktivitas seismik [8]. Pendekatan ini menegaskan pentingnya pengembangan model yang memanfaatkan data gempa historis serta menggabungkan observasi geofisika secara real-time [9]. Kerangka kerja ini menginspirasi inisiatif serupa di bidang terkait, seperti seismisitas yang dipicu, sehingga memperluas cakupan penelitian prediktabilitas gempa bumi. Sebagai contoh, evaluasi retrospektif prediksi gempa, sebagaimana terlihat dalam penelitian Woessner et al., menyoroti perlunya simulasi distribusi probabilitas tingkat seismisitas untuk menyempurnakan akurasi prediksi [10]. Penelitian-penelitian perbandingan semacam ini penting untuk memahami kinerja berbagai model prediksi dalam kondisi seismik yang berbeda.

Dalam penelitian seismik, beberapa fitur umum sangat penting untuk memahami dan memprediksi perilaku gempa bumi. Fitur-fitur ini meliputi magnitudo, kedalaman, lokasi, dan waktu, yang secara kolektif memberikan analisis komprehensif terhadap peristiwa seismik. Magnitudo merupakan parameter dasar yang mengukur energi yang dilepaskan selama gempa bumi. Biasanya diukur menggunakan skala magnitudo momen (M_w), yang memberikan representasi lebih akurat untuk peristiwa besar dibandingkan skala Richter. Penelitian menunjukkan bahwa distribusi magnitudo gempa dapat sangat mempengaruhi penilaian bahaya seismik dan model prediksi [11]. Misalnya, model gempa karakteristik yang sering digunakan dalam analisis bahaya seismik probabilistik (PSHA) sangat bergantung pada pemahaman distribusi magnitudo untuk memperkirakan potensi tingkat guncangan tanah [12]. Kedalaman gempa juga merupakan fitur penting, karena mempengaruhi intensitas

guncangan di permukaan. Gempa pada kedalaman dangkal cenderung menyebabkan guncangan yang lebih kuat dibandingkan gempa yang lebih dalam. Penelitian menunjukkan bahwa kedalaman gempa dapat memengaruhi distribusi spasial dari gempa susulan dan pendahulu, yang penting untuk memahami siklus seismik. Sebagai contoh, penelitian Heuret et al. menegaskan pentingnya kedalaman dalam karakterisasi zona seismogenik, di mana geometri dan perilaku mekanik dari antarmuka lempeng dapat memengaruhi terjadinya gempa [13]. Lokasi gempa, yang merujuk pada koordinat geografis, sangat penting untuk menilai distribusi spasial aktivitas seismik. Lokasi patahan dan seismisitas historis suatu wilayah merupakan faktor kritis dalam peramalan gempa. Model Fault Community (CFM) untuk California Selatan, misalnya, mengintegrasikan lokasi patahan untuk menyediakan sumber daya komprehensif bagi pemodelan sumber gempa dan penilaian bahaya seismik [14]. Waktu merupakan faktor penting yang terkait dengan pola temporal kejadian gempa. Memahami waktu kejadian gempa, termasuk interval antar gempa besar, penting untuk pengembangan model prediksi yang andal. Penelitian menunjukkan bahwa seismisitas sering menunjukkan perilaku pengelompokan, di mana kejadian kecil dapat mendahului gempa besar, yang menunjukkan akumulasi stres di kerak bumi.

Aplikasi pembelajaran mesin dalam geofisika, khususnya dalam prediksi peristiwa seismik, telah mengalami perkembangan pesat dalam beberapa tahun terakhir. Perkembangan ini didorong oleh ketersediaan data besar-besaran serta kebutuhan akan model prediksi yang lebih akurat untuk mengurangi risiko yang ditimbulkan oleh gempa bumi. Berbagai teknik pembelajaran mesin, seperti ANN, SVM, dan Random Forest, telah diterapkan untuk menganalisis data seismik dan meningkatkan akurasi prediksi. Salah satu keunggulan utama pembelajaran mesin dalam prediksi gempa adalah kemampuannya menangani data kompleks yang berdimensi tinggi. Tehseen et al. menyoroti penggunaan model hibrid yang menggabungkan algoritma genetika dengan ANN untuk mengoptimalkan prediksi peristiwa seismik, menunjukkan potensi ML untuk meningkatkan metode tradisional dalam seismologi [15]. Demikian pula, penelitian Lu terhadap data seismik multi-parameter menunjukkan bahwa penggabungan metode pembelajaran mesin yang berbeda dapat menghasilkan kinerja prediktif yang lebih baik dibandingkan pendekatan tunggal [16]. Jarah et al. juga menggunakan Random Forest untuk menganalisis data dari beberapa stasiun seismik, mengidentifikasi faktor-faktor kunci yang mendahului gempa bumi [17]. Selain itu, pembelajaran mesin telah terbukti efektif dalam aplikasi waktu nyata, seperti sistem peringatan dini gempa bumi (EEWS). Li et al. menekankan bahwa algoritma ML, setelah dilatih dengan data yang cukup, dapat beroperasi secara efisien dalam lingkungan waktu nyata, menjadikannya cocok untuk aplikasi yang memerlukan respons cepat terhadap aktivitas seismik [18]. Kemampuan ML untuk mengenali pola dan anomali dalam data seismik sangat penting untuk memberikan peringatan tepat waktu, seperti yang ditunjukkan oleh Corbi et al., yang menerapkan teknik ML untuk memprediksi gempa megathrust di zona subduksi [19].

Model pembelajaran mesin, seperti Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron, telah menunjukkan efektivitas yang signifikan dalam berbagai tugas prediktif di bidang geofisika, termasuk memprediksi magnitudo peristiwa seismik dan mengklasifikasikan aktivitas seismik. Model-model ini memanfaatkan kekuatan unik mereka untuk meningkatkan akurasi prediksi dan keandalan dalam aplikasi geofisika. Random Forest, sebagai metode pembelajaran ensemble yang membangun beberapa pohon keputusan, telah banyak digunakan karena ketangguhannya dan kemampuannya menangani data besar dengan dimensi tinggi. Efektivitas Random Forest dalam aplikasi seismik ditunjukkan oleh kemampuannya mengklasifikasikan peristiwa seismik dengan akurat. Izquierdo-Horna et al. menunjukkan integrasi Random Forest dalam penilaian risiko seismik, di mana Random Forest mampu memprediksi tingkat bahaya di berbagai skenario secara efektif [20]. Penggunaan Random Forest dalam mengklasifikasikan peristiwa seismik yang terkait dengan longsor juga disoroti oleh Provost et al., yang mencatat kemampuannya mengidentifikasi berbagai kelas seismisitas dengan tingkat akurasi yang tinggi [21]. Di sisi lain, Support Vector Regression unggul dalam tugas regresi, seperti memprediksi magnitudo peristiwa seismik. Asim et al. melaporkan bahwa Support Vector Regression, ketika dikombinasikan dengan jaringan saraf hibrid, secara signifikan meningkatkan akurasi prediksi magnitudo gempa di wilayah tertentu, menunjukkan potensinya dalam memperkuat model prediktif tradisional [22]. Kemampuan Support Vector Regression menemukan hyperplane yang paling sesuai dengan data memungkinkan model ini untuk secara efektif memodelkan hubungan kompleks dan non-linear dalam data seismik [23]. Penerapan Support Vector Regression dalam mengklasifikasikan peristiwa seismik juga dicatat oleh Sinambela et al., yang menggunakan teknik SVM untuk membedakan berbagai jenis aktivitas seismik, menunjukkan fleksibilitasnya dalam tugas regresi dan klasifikasi [24]. Model seperti XGBoost dan LightGBM, yang mengandalkan teknik boosting, serta Multi-Layer Perceptron dengan pendekatan jaringan saraf berlapis, juga berperan penting dalam menangkap hubungan non-linear dan pola kompleks pada data seismik, memberikan tambahan wawasan dalam analisis prediktif di bidang ini.

Pemodelan gempa bumi menghadapi berbagai tantangan, terutama akibat keterbatasan data, ketidaklinieran pada material geologi, serta variasi geografis. Faktor-faktor ini secara signifikan meningkatkan kompleksitas dalam memprediksi peristiwa seismik dengan akurat dan memahami dampaknya secara menyeluruh. Salah satu tantangan utama adalah keterbatasan data, yang memengaruhi akurasi model prediksi gempa. Ketersediaan dan

kualitas data seismik seringkali terbatas dan tidak merata di berbagai wilayah, sehingga memengaruhi keandalan hasil prediksi. Sebagai contoh, banyak model yang masih bergantung pada data gempa historis yang sering kali tidak komprehensif atau tersebar secara seragam, menyebabkan bias dalam pelatihan dan validasi model [25]. Selain itu, non-stasioneritas data gempa menyulitkan analisis statistik, karena metode tradisional sering kali mengasumsikan proses yang stasioner, sedangkan aktivitas seismik bersifat dinamis dan berubah-ubah [26]. Pertumbuhan pesat data seismik yang didorong oleh kemajuan teknologi pemantauan juga menambah kompleksitas ini, mengharuskan peneliti untuk mengembangkan metode yang mampu menangani data dalam skala besar, berdimensi tinggi, dan heterogen [27].

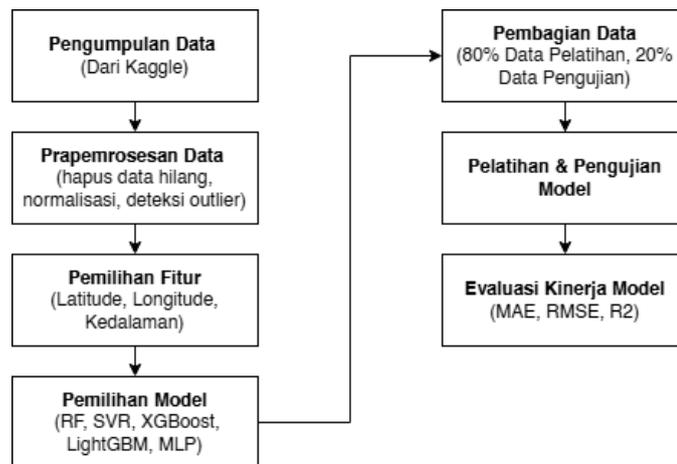
Ketidaklinieran adalah aspek penting lain yang mempersulit pemodelan gempa bumi. Perilaku material geologi di bawah tekanan sering kali tidak linier, terutama saat terjadi peristiwa seismik. Sebagai contoh, deposit tanah menunjukkan karakteristik tidak linier, seperti likuifaksi, yang dapat mengubah gerakan tanah dan respons struktural selama gempa [28]. Model linier tradisional mungkin gagal menangkap kompleksitas ini, sehingga menghasilkan penilaian risiko seismik yang kurang akurat [29]. Pendekatan pemodelan non-linier, seperti Non-Linear Time-History Analysis (NLTHA), telah dikembangkan untuk mengatasi tantangan ini, tetapi membutuhkan sumber daya komputasi yang besar dan karakterisasi material yang mendetail, yang kadang sulit diperoleh [30]. Variasi geografis menambah tantangan lain dalam pemodelan gempa bumi. Karakteristik geologi dan tektonik di berbagai wilayah dapat sangat bervariasi, yang memengaruhi bagaimana gelombang seismik merambat dan bagaimana struktur merespons getaran tanah. Misalnya, kedalaman hiposenter, kondisi permukaan, serta fitur geologi lokal dapat memengaruhi intensitas dan distribusi gelombang seismik [31]. Variabilitas geografis ini menuntut pengembangan model spesifik lokasi yang mampu mencerminkan kondisi lokal dengan akurat, namun proses ini sering kali memakan banyak sumber daya dan cukup kompleks [32]. Selain itu, interaksi antara berbagai proses fisik dan kimia di dalam Bumi menambah lapisan kompleksitas, karena interaksi ini bisa berbeda secara signifikan antara satu wilayah dengan wilayah lain.

Dalam bidang seismologi, penerapan teknik pembelajaran mesin terus berkembang, terutama untuk prediksi gempa dan estimasi magnitudo. Namun, terdapat kesenjangan yang signifikan dalam perbandingan langsung antara berbagai algoritma ML, seperti Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron, khususnya untuk prediksi magnitudo berdasarkan kedalaman dan lokasi. Meskipun penelitian sebelumnya telah menggunakan algoritma ini dalam berbagai konteks, hanya sedikit yang secara sistematis mengevaluasi kinerja mereka dalam konteks data seismik. Random Forest telah diakui karena kekuatannya dalam menangani hubungan non-linier dalam data, yang sangat penting dalam seismologi di mana pola semacam itu sering ditemukan [33]. Support Vector Regression, di sisi lain, menunjukkan kinerja yang baik pada dataset dengan ukuran terbatas, menjadikannya relevan untuk aplikasi prediktif khusus [34] [35]. Sementara itu, XGBoost dan LightGBM, yang berbasis pada teknik boosting, menawarkan kemampuan dalam menangkap pola-pola kompleks pada data, sedangkan Multi-Layer Perceptron dengan pendekatan jaringan saraf berlapis memberikan fleksibilitas tambahan dalam menangani hubungan non-linear. Namun, penelitian komparatif yang berfokus pada efektivitas algoritma-algoritma ini dalam memprediksi magnitudo gempa berdasarkan fitur geologis seperti kedalaman, lintang, dan bujur masih jarang dilakukan.

Penelitian ini bertujuan untuk memprediksi magnitudo gempa bumi berdasarkan data kedalaman, lintang, dan bujur dengan membandingkan kinerja lima algoritma pembelajaran mesin: Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron. Algoritma-algoritma ini dipilih karena kemampuannya dalam menangani hubungan non-linear dan kompleksitas data geofisika, yang menjadi tantangan utama dalam prediksi seismik. Penelitian ini mengevaluasi efektivitas setiap algoritma melalui perbandingan kinerja prediksi berdasarkan metrik evaluasi yang relevan seperti MAE, RMSE, dan R^2 . Dengan pendekatan komparatif ini, diharapkan dapat memberikan kontribusi yang signifikan terhadap pengembangan model prediksi gempa yang lebih akurat dan efektif untuk meningkatkan kesiapsiagaan dan mitigasi risiko di wilayah yang rawan gempa.

2. METODE PENELITIAN

Penelitian ini dimulai dengan pengumpulan data gempa dari Kaggle. Data kemudian melalui tahapan pra-pemrosesan, termasuk penanganan data hilang, normalisasi menggunakan Min-Max Scaler, dan deteksi outlier dengan Z-score. Fitur Latitude, Longitude, dan Kedalaman dipilih sebagai prediktor untuk memprediksi Magnitudo gempa. Selanjutnya, lima model pembelajaran mesin digunakan untuk memprediksi magnitudo gempa. Data dibagi menjadi data pelatihan (80%) dan data pengujian (20%). Pada Gambar 1, ditampilkan alur proses penelitian yang meliputi tahapan mulai dari pengumpulan data, pra-pemrosesan data, pemilihan fitur, pemodelan dengan lima algoritma pembelajaran mesin, hingga evaluasi kinerja model. Gambar ini menggambarkan secara jelas urutan langkah-langkah yang diambil dalam penelitian untuk memprediksi magnitudo gempa menggunakan data kedalaman dan geolokasi.



Gambar 1. Diagram Alir Metode Penelitian

2.1. Deskripsi Dataset

Penelitian ini menggunakan dataset yang berasal dari Kaggle yang mencakup rentang waktu dari tahun 2012 hingga 2024 dengan jumlah data sebanyak 3885 records. Dataset ini memuat sejumlah fitur utama, termasuk Tanggal, Latitude, Longitude, Kedalaman, dan Magnitudo dari kejadian gempa bumi yang tercatat. Data ini dipilih karena mencerminkan distribusi geografis dan aktivitas seismik di wilayah tertentu, sehingga relevan untuk analisis prediksi magnitudo gempa. Informasi geolokasi seperti latitude dan longitude memberikan gambaran spasial, sementara kedalaman merupakan parameter penting yang berkaitan dengan intensitas dan dampak gempa di permukaan. Sebelum analisis lebih lanjut, data melalui sejumlah tahap pra-pemrosesan untuk memastikan keakuratannya. Langkah pertama adalah penanganan data yang hilang, di mana setiap baris dengan data kosong dihapus guna menghindari bias dalam model prediksi. Selanjutnya, normalisasi fitur dilakukan menggunakan metode Min-Max Scaler untuk mengubah nilai-nilai Latitude, Longitude, Kedalaman, dan Magnitudo ke dalam rentang [0,1]. Hal ini bertujuan agar fitur memiliki skala yang seragam, sehingga model pembelajaran mesin dapat memproses data secara optimal. Selain itu, deteksi outlier dilakukan menggunakan metode Z-score dengan batas lebih dari tiga standar deviasi dari rata-rata. Pendekatan ini bertujuan untuk mengidentifikasi dan menangani data yang berpotensi menjadi anomali, yang dapat mempengaruhi keakuratan prediksi.

2.2. Pemilihan Fitur

Dalam penelitian ini, fitur Latitude, Longitude, dan Kedalaman dipilih sebagai prediktor untuk memprediksi magnitudo gempa bumi. Pemilihan fitur ini didasarkan pada pengaruh signifikan dari parameter geografis dan geologis terhadap kekuatan gempa. Latitude dan Longitude merepresentasikan lokasi geografis episentrum gempa, yang dapat memengaruhi penyebaran gelombang seismik serta dampak kerusakan di permukaan. Sementara itu, Kedalaman mencerminkan seberapa dalam gempa terjadi di bawah permukaan bumi, yang berperan penting dalam menentukan intensitas getaran di permukaan. Gempa dangkal cenderung menyebabkan kerusakan yang lebih besar dibandingkan gempa yang terjadi lebih dalam, meskipun dengan magnitudo yang sama. Dengan demikian, kombinasi ketiga fitur ini diharapkan dapat memberikan gambaran prediktif yang lebih akurat mengenai besaran magnitudo gempa.

2.3. Pemilihan Model

Pemodelan dalam penelitian ini menggunakan lima algoritma pembelajaran mesin, yaitu Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron. Random Forest merupakan metode ensemble yang membangun beberapa pohon keputusan dan menggabungkan prediksinya untuk menghasilkan hasil yang lebih akurat dan stabil. Algoritma ini sangat efektif dalam menangani hubungan non-linear antar fitur serta mencegah overfitting dengan cara merata-ratakan hasil dari banyak pohon keputusan. Kemampuan Random Forest dalam mengolah data berdimensi tinggi menjadikannya sangat sesuai untuk analisis data seismik yang kompleks.

Support Vector Regression adalah algoritma regresi yang menggunakan fungsi kernel untuk memodelkan hubungan non-linear antara fitur dan variabel target. Support Vector Regression bekerja dengan memproyeksikan data ke ruang berdimensi lebih tinggi untuk menemukan hyperplane yang meminimalkan

kesalahan prediksi. Kekuatan Support Vector Regression dalam menangani data yang tidak dapat dipisahkan secara linear membuatnya cocok untuk tugas prediksi gempa, terutama ketika hubungan antar variabel bersifat kompleks dan sulit diprediksi dengan metode linear konvensional. Sementara itu, XGBoost dan LightGBM, yang berbasis pada metode boosting, menawarkan efisiensi tinggi dalam menangkap pola-pola non-linear dan menghadirkan interpretabilitas tambahan melalui fitur penting. Multi-Layer Perceptron dengan jaringan saraf berlapis memberikan fleksibilitas dalam menangani data kompleks melalui pendekatan non-linear yang canggih.

2.4. Pembagian Data Pelatihan dan Pengujian

Data dibagi menjadi dua set, yaitu data pelatihan (80%) dan data pengujian (20%). Pembagian ini dilakukan secara acak untuk memastikan bahwa model dapat dilatih dengan data yang representatif serta diuji keakuratannya pada data yang tidak terlihat sebelumnya. Metode ini bertujuan untuk mengevaluasi kemampuan generalisasi model dalam memprediksi magnitudo gempa berdasarkan fitur yang telah dipilih.

2.5. Metode Evaluasi

Evaluasi kinerja model dilakukan menggunakan beberapa metrik, yaitu MAE, RMSE, dan R^2 score. MAE mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai sebenarnya, memberikan gambaran seberapa besar kesalahan rata-rata prediksi model. RMSE menghitung akar dari rata-rata kesalahan kuadrat, memberikan penalti lebih besar pada kesalahan yang lebih besar, sehingga memberikan indikasi sejauh mana prediksi menyimpang dari nilai sebenarnya. R^2 score menunjukkan proporsi variansi dalam data target yang dapat dijelaskan oleh model, dengan nilai mendekati 1 menunjukkan performa prediktif yang lebih baik.

2.6. Detail Implementasi

Implementasi penelitian ini menggunakan bahasa pemrograman Python dengan memanfaatkan berbagai pustaka pembelajaran mesin dan analisis data. Pustaka utama yang digunakan meliputi scikit-learn untuk pemodelan algoritma Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron, serta pandas dan NumPy untuk pengolahan data. Visualisasi data dilakukan dengan bantuan pustaka Matplotlib dan Seaborn untuk memperoleh pemahaman yang lebih dalam mengenai distribusi dan korelasi antar fitur. Pemilihan scikit-learn didasarkan pada fleksibilitasnya dan kelengkapannya dalam menyediakan alat untuk pemodelan, pelatihan, evaluasi, serta tuning parameter model. Untuk memastikan kinerja algoritma optimal, dilakukan proses hyperparameter tuning pada setiap model. Pada Random Forest, parameter utama yang disesuaikan meliputi jumlah pohon ($n_{\text{estimators}}$) yang digunakan untuk membangun hutan, kedalaman maksimum pohon (max_depth) untuk mengontrol kompleksitas model, minimum samples split yang menentukan jumlah minimum sampel yang dibutuhkan untuk membagi simpul, dan minimum samples leaf yang menetapkan jumlah minimum sampel pada simpul daun. Parameter ini dituning menggunakan teknik grid search untuk menemukan kombinasi terbaik yang mengurangi risiko overfitting.

Pada Support Vector Regression, kernel RBF (Radial Basis Function) digunakan untuk menangani hubungan non-linear antara data. Parameter yang disesuaikan meliputi parameter regulasi C, yang mengontrol tingkat penalti terhadap kesalahan, gamma yang mempengaruhi seberapa jauh pengaruh satu sampel data meluas, dan epsilon yang menentukan margin kesalahan yang dapat diterima tanpa penalti dalam model. Grid search digunakan untuk mengoptimalkan parameter ini agar menghasilkan performa prediktif yang maksimal. XGBoost menggunakan parameter seperti learning rate yang mengontrol kecepatan pembelajaran, $n_{\text{estimators}}$ yang menunjukkan jumlah pohon yang dibangun, dan max_depth untuk mengontrol kedalaman pohon. Selain itu, subsample menentukan fraksi sampel yang dipilih untuk membangun setiap pohon, dan colsample_bytree mengatur proporsi fitur yang digunakan dalam setiap iterasi. LightGBM juga menggunakan parameter serupa seperti learning rate, num_leaves yang mengontrol jumlah daun maksimum dalam setiap pohon, max_depth , dan min_child_samples yang mengatur jumlah minimum sampel di setiap simpul daun untuk menghindari overfitting.

Untuk Multi-Layer Perceptron, parameter yang disesuaikan meliputi jumlah layer tersembunyi dan jumlah neuron pada setiap layer, fungsi aktivasi yang digunakan (misalnya ReLU), nilai learning rate untuk optimisasi, dan jumlah iterasi maksimum (max_iter) untuk pelatihan. Hyperparameter tuning dilakukan untuk memastikan kinerja model Multi-Layer Perceptron yang optimal, dengan penyesuaian pada parameter seperti batch_size untuk mempercepat pelatihan dan regulasi (L2) untuk mengurangi risiko overfitting. Proses tuning pada semua model dilakukan dengan pembagian data menggunakan k-fold cross-validation guna memastikan hasil yang robust dan memiliki generalisasi yang baik pada data uji. Pengaturan parameter yang optimal diharapkan mampu meningkatkan performa prediktif dan akurasi model dalam memprediksi magnitudo gempa.

3. HASIL DAN PEMBAHASAN

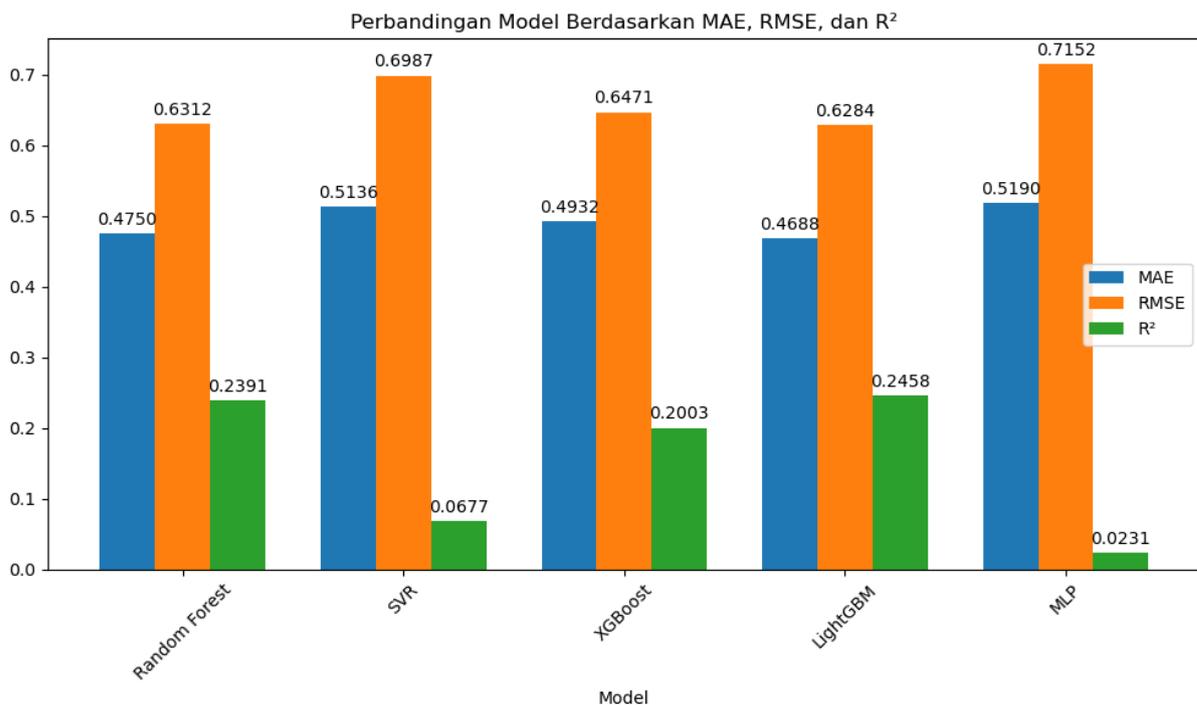
3.1. Kinerja Model

Kinerja model pembelajaran mesin dievaluasi menggunakan tiga metrik utama, yaitu Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan R² score, untuk membandingkan performa masing-masing algoritma dalam memprediksi magnitudo gempa, seperti ditampilkan pada Tabel 1. Random Forest menunjukkan hasil yang baik dengan nilai MAE sebesar 0.4750, RMSE sebesar 0.6312, dan R² sebesar 0.2391, yang menunjukkan bahwa model ini mampu menangkap sekitar 23,91% dari variabilitas dalam data target. Kemampuan Random Forest untuk menangani data berdimensi tinggi dan hubungan non-linear menjadikannya efektif dalam tugas prediksi gempa ini. Sebaliknya, Support Vector Regression mencatatkan kinerja yang lebih rendah, dengan MAE sebesar 0.5136, RMSE sebesar 0.6987, dan R² sebesar 0.0677. Dengan hanya mampu menjelaskan sekitar 6,77% dari variabilitas target, Support Vector Regression menunjukkan bahwa meskipun efektif dalam beberapa skenario, algoritma ini kurang berhasil dalam memodelkan data gempa yang kompleks.

Tabel 1. Perbandingan Performa Model

Model	MAE	RMSE	R ²
Random Forest	0.475	0.6312	0.2391
Support Vector Regression	0.5136	0.6987	0.0677
XGBoost	0.4932	0.6471	0.2003
LightGBM	0.4688	0.6284	0.2458
Multi-Layer Perceptron	0.519	0.7152	0.0231

Selain itu, model XGBoost dan LightGBM juga diuji untuk membandingkan kinerjanya. LightGBM memberikan hasil yang sedikit lebih baik daripada Random Forest dengan MAE sebesar 0.4688, RMSE sebesar 0.6284, dan R² sebesar 0.2458, menunjukkan bahwa model ini memiliki keunggulan dalam menangani data seismik dan mampu memberikan prediksi yang akurat. XGBoost menunjukkan kinerja yang kompetitif dengan nilai MAE sebesar 0.4932, RMSE sebesar 0.6471, dan R² sebesar 0.2003. Sementara itu, Multi-Layer Perceptron mencatatkan performa terburuk dengan MAE sebesar 0.5190, RMSE sebesar 0.7152, dan R² sebesar 0.0231. Hal ini menunjukkan bahwa Multi-Layer Perceptron kurang efektif dalam menangkap hubungan kompleks antara fitur seperti kedalaman, lintang, dan bujur dengan magnitudo gempa. Hasil ini menunjukkan bahwa pemilihan algoritma yang tepat sangat penting dalam meningkatkan akurasi prediksi gempa, sebagaimana divisualisasikan dalam grafik perbandingan performa model pada Gambar 2.



Gambar 2. Grafik Perbandingan Performa Model

3.2. Interpretasi Hasil

Hasil perbandingan kinerja lima model pembelajaran mesin dalam memprediksi magnitudo gempa menunjukkan keunggulan tertentu pada model Random Forest dan LightGBM, sementara model lainnya menunjukkan performa yang relatif lebih rendah. Random Forest mencatatkan nilai MAE sebesar 0.475, RMSE sebesar 0.6312, dan R^2 sebesar 0.2391. Hasil ini menunjukkan bahwa Random Forest mampu menjelaskan sekitar 23,91% dari variansi dalam data target, menandakan kemampuan yang cukup baik dalam menangkap hubungan kompleks antara variabel prediktor seperti kedalaman, lintang, dan bujur dengan magnitudo gempa. Keunggulan Random Forest terletak pada penggunaan metode ensemble yang menggabungkan banyak pohon keputusan. Proses ini memungkinkan model untuk menangani hubungan non-linear antar fitur dengan lebih efektif dan meningkatkan akurasi prediksi secara keseluruhan. Selain itu, mekanisme pembobotan pohon dan proses penggabungan prediksi membantu mengurangi overfitting, menjadikannya lebih robust terhadap variasi data. Sebaliknya, Support Vector Regression menunjukkan kinerja yang kurang optimal dengan nilai MAE sebesar 0.5136, RMSE sebesar 0.6987, dan R^2 sebesar 0.0677. Nilai R^2 yang rendah menunjukkan bahwa model ini hanya mampu menjelaskan sekitar 6,77% dari variabilitas data target. Hal ini menunjukkan bahwa Support Vector Regression memiliki keterbatasan dalam menangkap pola kompleks yang muncul dalam data seismik. Salah satu faktor yang berkontribusi pada keterbatasan ini adalah ketergantungan Support Vector Regression pada pemilihan kernel yang tepat, di mana kernel RBF yang digunakan mungkin tidak cukup fleksibel untuk memodelkan seluruh kompleksitas hubungan antar variabel. Selain itu, kinerja Support Vector Regression cenderung menurun pada dataset dengan dimensi tinggi atau data yang memiliki distribusi kompleks, seperti yang terdapat dalam data seismik.

Model XGBoost, dengan MAE sebesar 0.4932, RMSE sebesar 0.6471, dan R^2 sebesar 0.2003, menunjukkan kinerja yang cukup baik meskipun sedikit di bawah Random Forest dan LightGBM. XGBoost memanfaatkan boosting berbasis pohon keputusan yang secara iteratif meningkatkan kesalahan prediksi dari model sebelumnya. Meskipun hasilnya tidak sebaik LightGBM, XGBoost tetap memberikan prediksi yang kompetitif dan menunjukkan kemampuannya dalam menangani data kompleks dengan interaksi non-linear antar fitur. Keunggulan XGBoost terletak pada efisiensinya dalam waktu pelatihan serta kemampuannya untuk mengatasi overfitting melalui regularisasi yang lebih baik. LightGBM mencatatkan performa terbaik dengan nilai MAE sebesar 0.4688, RMSE sebesar 0.6284, dan R^2 sebesar 0.2458. Model ini menunjukkan kemampuan prediksi yang sangat baik berkat struktur algoritmanya yang dioptimalkan untuk menangani dataset yang besar dengan lebih cepat dibandingkan model lainnya. LightGBM memanfaatkan pendekatan pembagian pohon berbasis histogram, yang memungkinkan penanganan data skala besar dengan waktu komputasi yang lebih efisien. Kemampuan LightGBM untuk menangani outlier dan interaksi fitur yang kompleks menjadikannya sangat cocok untuk prediksi gempa berdasarkan data kedalaman dan geolokasi.

Sebaliknya, Multi-Layer Perceptron menunjukkan performa terendah dengan nilai MAE sebesar 0.519, RMSE sebesar 0.7152, dan R^2 sebesar 0.0231. Hal ini menunjukkan bahwa Multi-Layer Perceptron mengalami kesulitan dalam memodelkan hubungan antar fitur dalam data seismik. Kemungkinan penyebabnya termasuk arsitektur jaringan yang tidak cukup dalam atau kurang optimalnya parameter pelatihan. Selain itu, sifat data seismik yang memiliki hubungan non-linear kompleks mungkin tidak sepenuhnya dapat ditangkap oleh Multi-Layer Perceptron tanpa penyesuaian arsitektur yang signifikan. Hal ini menunjukkan bahwa model deep learning seperti Multi-Layer Perceptron membutuhkan tuning yang lebih ekstensif dan penanganan data yang lebih baik untuk memberikan prediksi yang akurat. Secara keseluruhan, hasil ini menunjukkan bahwa model ensemble seperti Random Forest dan LightGBM lebih unggul dalam menangani kompleksitas data seismik dibandingkan dengan Support Vector Regression dan Multi-Layer Perceptron. XGBoost memberikan performa yang kompetitif, sementara Multi-Layer Perceptron menunjukkan keterbatasan dalam menangani data seismik yang kompleks tanpa penyesuaian parameter yang tepat. Temuan ini menyoroti pentingnya pemilihan algoritma yang sesuai berdasarkan karakteristik data dan kebutuhan prediksi yang dihadapi.

3.3. Kekuatan dan Keterbatasan Model

Setiap model pembelajaran mesin yang digunakan dalam penelitian ini memiliki kekuatan dan keterbatasan dalam memprediksi magnitudo gempa berdasarkan data kedalaman dan geolokasi. Model Random Forest menunjukkan kekuatan utamanya dalam menangani hubungan non-linear antar fitur dan dataset dengan dimensi tinggi. Kemampuan model ini untuk membangun beberapa pohon keputusan dan melakukan proses voting atau averaging dari hasil prediksi setiap pohon memberikan kestabilan dan keakuratan prediksi yang lebih baik. Random Forest juga memiliki ketahanan terhadap overfitting berkat proses ensemble, yang membuatnya lebih mampu menggeneralisasi data baru. Namun, model ini cenderung membutuhkan waktu komputasi yang lebih lama dan memori yang lebih besar seiring bertambahnya jumlah pohon yang digunakan, yang dapat menjadi tantangan saat digunakan pada dataset yang sangat besar. Support Vector Regression memiliki kekuatan dalam

memodelkan hubungan non-linear melalui penggunaan fungsi kernel, seperti kernel Radial Basis Function (RBF). Support Vector Regression bekerja dengan memproyeksikan data ke dimensi yang lebih tinggi untuk menemukan hyperplane yang meminimalkan kesalahan prediksi. Kelebihan Support Vector Regression terletak pada kemampuannya menangani dataset yang lebih kecil dengan baik dan menghasilkan prediksi yang cukup presisi dalam kasus tertentu. Meskipun demikian, Support Vector Regression memiliki keterbatasan dalam hal skalabilitas pada dataset yang besar, di mana proses komputasi dapat menjadi sangat lambat dan memakan banyak sumber daya. Selain itu, pemilihan parameter yang optimal, seperti parameter regulasi (C) dan parameter kernel (γ), sangat mempengaruhi kinerja model, sehingga memerlukan tuning yang hati-hati.

XGBoost menawarkan kekuatan utama dalam hal efisiensi waktu pelatihan dan penanganan hubungan non-linear yang kompleks melalui pendekatan boosting berbasis pohon. Model ini sangat efektif dalam menangani outlier dan data skala besar dengan tingkat akurasi yang baik. Namun, XGBoost dapat menghadapi risiko overfitting jika parameter tuning tidak dilakukan dengan benar. LightGBM, sebagai model yang dioptimalkan untuk efisiensi dan kecepatan, menunjukkan kekuatan serupa dengan XGBoost, tetapi lebih cepat dalam menangani dataset besar berkat pendekatan pembagian pohon berbasis histogram. Kekurangan utama LightGBM adalah sensitif terhadap parameter hyperparameter, yang memerlukan perhatian khusus dalam tuning. Multi-Layer Perceptron memiliki keunggulan dalam menangani data dengan hubungan non-linear yang kompleks melalui jaringan lapisan tersembunyi. Namun, model ini menunjukkan keterbatasan dalam memprediksi magnitudo gempa dalam penelitian ini, kemungkinan disebabkan oleh arsitektur jaringan yang kurang optimal dan sensitivitas terhadap overfitting. Multi-Layer Perceptron membutuhkan data dalam jumlah besar dan tuning parameter yang ekstensif agar dapat memberikan hasil yang akurat, yang dapat menjadi tantangan dalam konteks data seismik dengan karakteristik unik. Hal ini menunjukkan bahwa Multi-Layer Perceptron memerlukan penyesuaian arsitektur jaringan dan pengaturan parameter yang lebih cermat untuk menangkap hubungan yang ada dalam data seismik secara efektif.

3.4. Implikasi untuk Seismologi berbasis Teknologi

Temuan dari penelitian ini memiliki potensi kontribusi yang signifikan bagi bidang prediksi gempa bumi melalui pendekatan berbasis teknologi. Dengan menunjukkan bahwa model Random Forest memberikan hasil yang lebih unggul dalam memprediksi magnitudo gempa berdasarkan data kedalaman dan geolokasi, pendekatan ini dapat memperkuat keakuratan model prediksi yang ada. Ketepatan prediksi magnitudo gempa sangat penting dalam mendukung perencanaan mitigasi risiko dan langkah-langkah kesiapsiagaan yang berbasis data, yang pada akhirnya dapat menyelamatkan banyak nyawa serta mengurangi dampak kerugian ekonomi. Model pembelajaran mesin seperti ini mampu menangkap hubungan kompleks antara variabel kedalaman, lokasi, dan magnitudo, memberikan wawasan lebih dalam tentang pola perilaku gempa bumi yang kompleks, sehingga memungkinkan para ahli seismologi untuk menganalisis tren dan karakteristik seismik dengan pendekatan berbasis data yang lebih mendalam.

Aplikasi dunia nyata dari penerapan model pembelajaran mesin ini mencakup peningkatan pada sistem peringatan dini gempa bumi dan evaluasi risiko seismik yang lebih canggih dan komprehensif. Dengan mengintegrasikan model seperti Random Forest ke dalam sistem peringatan dini berbasis teknologi, kecepatan dan akurasi respons dapat ditingkatkan, memberikan masyarakat dan otoritas lebih banyak waktu untuk mempersiapkan diri sebelum dampak gempa dirasakan. Selain itu, penerapan model ini dalam penilaian risiko seismik juga memungkinkan identifikasi area berisiko tinggi dengan lebih akurat, mendukung pengambilan keputusan dalam perencanaan tata ruang berbasis data, pembangunan infrastruktur yang tahan gempa, serta kebijakan mitigasi bencana yang lebih terstruktur dan efektif.

3.5. Tantangan dan Penelitian Selanjutnya

Penelitian ini menghadapi beberapa keterbatasan, salah satunya adalah ukuran dataset yang relatif terbatas. Keterbatasan ini dapat memengaruhi generalisasi hasil prediksi model ketika diterapkan pada data gempa di lokasi berbeda. Selain itu, variabel yang digunakan, yakni kedalaman, latitude, dan longitude, meskipun berkontribusi secara signifikan terhadap prediksi magnitudo gempa, tidak mencakup beberapa faktor penting lainnya seperti data tentang garis patahan atau karakteristik geologi lainnya. Kompleksitas fenomena gempa bumi yang sangat bergantung pada banyak variabel fisik dan mekanis menambah tantangan dalam mengembangkan model prediksi yang akurat dan andal.

Untuk penelitian di masa depan, penambahan fitur-fitur yang lebih komprehensif, seperti tipe patahan, struktur geologi setempat, serta data historis tambahan, dapat membantu meningkatkan akurasi model. Selain itu, membandingkan kinerja algoritma pembelajaran mesin lainnya, seperti model deep learning atau metode ensemble lainnya, dapat memberikan wawasan yang lebih mendalam dan mungkin menghasilkan pendekatan yang lebih optimal untuk prediksi gempa bumi. Penggunaan data geofisika yang lebih detail serta kolaborasi

multidisiplin antara ahli seismologi, data scientist, dan ahli geologi diharapkan dapat mendukung pengembangan model prediksi yang lebih canggih dan relevan dengan kebutuhan mitigasi bencana gempa di masa depan.

4. KESIMPULAN

Penelitian ini membandingkan kinerja berbagai model pembelajaran mesin, yaitu Random Forest, Support Vector Regression, XGBoost, LightGBM, dan Multi-Layer Perceptron dalam memprediksi magnitudo gempa berdasarkan data kedalaman dan geolokasi. Hasil menunjukkan bahwa LightGBM memberikan kinerja terbaik dengan nilai MAE dan Root RMSE terendah serta nilai R^2 yang lebih tinggi dibandingkan model lainnya, menandakan kemampuan prediksi yang akurat dan stabil dalam menjelaskan variabilitas data target. Random Forest juga menunjukkan kinerja yang sangat baik dengan nilai MAE, RMSE, dan R^2 yang sebanding, menjadikannya pilihan yang kuat untuk memodelkan data yang kompleks dan non-linear dalam konteks prediksi gempa. Sebaliknya, Support Vector Regression menunjukkan hasil yang kurang optimal, dengan nilai MAE dan RMSE yang lebih tinggi serta nilai R^2 yang lebih rendah, mengindikasikan bahwa model ini kurang efektif dalam menangkap variabilitas data target dibandingkan model ensemble seperti Random Forest dan LightGBM. XGBoost memberikan hasil yang cukup baik, meskipun kinerjanya sedikit di bawah LightGBM dan Random Forest. Sementara itu, Multi-Layer Perceptron mencatatkan kinerja terendah, dengan nilai MAE dan RMSE tertinggi serta nilai R^2 yang sangat rendah, yang menunjukkan kesulitan model ini dalam menangkap hubungan kompleks antar fitur dalam data seismik. Secara keseluruhan, penelitian ini menunjukkan bahwa model ensemble seperti Random Forest dan LightGBM lebih unggul dalam memprediksi magnitudo gempa, khususnya dalam konteks data geofisika yang kompleks dan non-linear. Hasil ini menekankan pentingnya pemilihan algoritma yang sesuai untuk aplikasi seismologi guna meningkatkan akurasi prediksi dan mendukung upaya mitigasi risiko bencana. Diharapkan temuan ini dapat menjadi dasar untuk pengembangan model prediksi yang lebih canggih dan aplikatif dalam upaya pengurangan dampak gempa bumi di masa depan. Penelitian ini menyoroti pentingnya perbandingan antara algoritma pembelajaran mesin dalam prediksi gempa bumi, memberikan wawasan yang lebih dalam mengenai hubungan antara data geolokasi dan kedalaman dengan magnitudo gempa. Perbandingan ini berkontribusi pada peningkatan model prediksi yang dapat membantu mitigasi risiko bencana gempa dan meningkatkan pemahaman mengenai perilaku seismik di berbagai wilayah. Implikasi praktis dari penelitian ini dapat diterapkan dalam peningkatan penilaian risiko seismik, kesiapsiagaan bencana, serta sistem peringatan dini di wilayah rawan gempa seperti Chile dan kawasan seismik lainnya. Sebagai penutup, potensi penerapan pembelajaran mesin dalam bidang seismologi menunjukkan arah baru yang menjanjikan dalam meningkatkan akurasi prediksi gempa, namun tetap membutuhkan penyempurnaan berkelanjutan agar relevan dan dapat diterapkan di dunia nyata.

DAFTAR PUSTAKA

- [1] Z. Zhao *dkk.*, “A Seismic Emergency Performance Optimization Model for Infrastructure Systems Under Demand Differences: A Case Study in China,” *Earthq. Eng. Resil.*, vol. 1, no. 2, hlm. 196–210, 2022, doi: 10.1002/eer2.22.
- [2] J. Chu, Q. Zhang, W. Ai, dan H. Yu, “A Hybrid Intelligent Model for Urban Seismic Risk Assessment From the Perspective of Possibility and Vulnerability Based on Particle Swarm Optimization,” *Sci. Program.*, vol. 2021, hlm. 1–16, 2021, doi: 10.1155/2021/2218044.
- [3] Z. A. Lamjiry dan R. Gifford, “Earthquake Threat! Understanding the Intention to Prepare for the Big One,” *Risk Anal.*, vol. 42, no. 3, hlm. 487–505, 2021, doi: 10.1111/risa.13775.
- [4] R. Ternero, “Analysis of Pedestrian Behavior for the Optimization of Evacuation Plans in Tall Buildings: Case Study Santiago, Chile,” *Buildings*, vol. 13, no. 12, hlm. 2907, 2023, doi: 10.3390/buildings13122907.
- [5] V. Scapini, “Disaster, Infrastructure Damage, and Health,” *Int. J. Saf. Secur. Eng.*, vol. 10, no. 2, hlm. 219–225, 2020, doi: 10.18280/ijss.100208.
- [6] D. B. d. Freitas, G. S. França, T. M. Scherrer, C. S. Vilar, dan R. Silva, “Investigating the Signatures of Long-Range Persistence in Seismic Sequences Along Circum-Pacific Subduction Zones,” *Braz. J. Geophys.*, vol. 37, no. 4, hlm. 409, 2019, doi: 10.22564/rbfg.v37i4.2018.
- [7] V. Scapini dan E. Zuñiga, “A Markov Chain Approach to Model Reconstruction,” *Int. J. Comput. Methods Exp. Meas.*, vol. 8, no. 4, hlm. 316–325, 2020, doi: 10.2495/cm-m-v8-n4-316-325.
- [8] T. Wang, J. Zhuang, dan T. Kato, “Assessing the Potential Improvement in Short-term Earthquake Forecasts From Incorporation of GPS Data,” *Geophys. Res. Lett.*, vol. 40, no. 11, hlm. 2631–2635, 2013, doi: 10.1002/grl.50554.

- [9] D. Schorlemmer *dkk.*, “The Collaboratory for the Study of Earthquake Predictability: Achievements and Priorities,” *Seismol. Res. Lett.*, vol. 89, no. 4, hlm. 1305–1313, 2018, doi: 10.1785/0220180053.
- [10] J. Woessner *dkk.*, “A Retrospective Comparative Forecast Test on the 1992 Landers Sequence,” *J. Geophys. Res. Atmospheres*, vol. 116, no. B5, 2011, doi: 10.1029/2010jb007846.
- [11] M. T. Page, D. Alderson, dan J. Doyle, “The Magnitude Distribution of Earthquakes Near Southern California Faults,” *J. Geophys. Res. Atmospheres*, vol. 116, no. B12, 2011, doi: 10.1029/2010jb007933.
- [12] A. Akinci, “HAZGRIDX: Earthquake Forecasting Model for $ML \geq 5.0$ Earthquakes in Italy Based on Spatially Smoothed Seismicity,” *Ann. Geophys.*, vol. 53, no. 3, 2010, doi: 10.4401/ag-4811.
- [13] A. Heuret, S. Lallemand, F. Funiciello, C. Piromallo, dan C. Faccenna, “Physical Characteristics of Subduction Interface Type Seismogenic Zones Revisited,” *Geochem. Geophys. Geosystems*, vol. 12, no. 1, hlm. n/a-n/a, 2011, doi: 10.1029/2010gc003230.
- [14] A. Plesch *dkk.*, “Community Fault Model (CFM) for Southern California,” *Bull. Seismol. Soc. Am.*, vol. 97, no. 6, hlm. 1793–1802, 2007, doi: 10.1785/0120050211.
- [15] R. Tehseen, M. S. Farooq, dan A. Abid, “Earthquake Prediction Using Expert Systems: A Systematic Mapping Study,” *Sustainability*, vol. 12, no. 6, hlm. 2420, 2020, doi: 10.3390/su12062420.
- [16] X. Lu, “Machine Learning-Based Precursor Detection Using Seismic Multi-Parameter Data,” *Appl. Sci.*, vol. 14, no. 6, hlm. 2401, 2024, doi: 10.3390/app14062401.
- [17] N. B. Jarah, A. H. H. Alasadi, dan K. M. Hashim, “Earthquake Prediction Technique: A Comparative Study,” *Iaes Int. J. Artif. Intell. Ij-Ai*, vol. 12, no. 3, hlm. 1026, 2023, doi: 10.11591/ijai.v12.i3.pp1026-1032.
- [18] Z. Li, M. Meier, E. Hauksson, Z. Zhan, dan J. Andrews, “Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning,” *Geophys. Res. Lett.*, vol. 45, no. 10, hlm. 4773–4779, 2018, doi: 10.1029/2018gl077870.
- [19] F. Corbi, J. Bedford, L. Sandri, F. Funiciello, A. Gualandi, dan M. Rosenau, “Predicting Imminence of Analog Megathrust Earthquakes With Machine Learning: Implications for Monitoring Subduction Zones,” *Geophys. Res. Lett.*, vol. 47, no. 7, 2020, doi: 10.1029/2019gl086615.
- [20] L. Izquierdo-Horna, J. Zevallos, dan Y. YEPEZ, “An Integrated Approach to Seismic Risk Assessment Using Random Forest and Hierarchical Analysis: Pisco, Peru,” *Heliyon*, vol. 8, no. 10, hlm. e10926, 2022, doi: 10.1016/j.heliyon.2022.e10926.
- [21] F. Provost, C. Hibert, dan J. Malet, “Automatic Classification of Endogenous Landslide Seismicity Using the Random Forest Supervised Classifier,” *Geophys. Res. Lett.*, vol. 44, no. 1, hlm. 113–120, 2017, doi: 10.1002/2016gl070709.
- [22] K. M. Asim, A. Idris, T. Iqbal, dan F. Martínez-Álvarez, “Earthquake Prediction Model Using Support Vector Regressor and Hybrid Neural Networks,” *Plos One*, vol. 13, no. 7, hlm. e0199004, 2018, doi: 10.1371/journal.pone.0199004.
- [23] J. Yin, “A Semismooth Newton Method for Support Vector Classification and Regression,” *Comput. Optim. Appl.*, vol. 73, no. 2, hlm. 477–508, 2019, doi: 10.1007/s10589-019-00075-z.
- [24] M. Sinambela, M. Situmorang, K. Tarigan, S. Humaidi, dan M. Sirait, “Waveforms Classification of Northern Sumatera Earthquakes for New Mini Region Stations Using Support Vector Machine,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 11, no. 2, hlm. 489–494, 2021, doi: 10.18517/ijaseit.11.2.12503.
- [25] “Technical Bottlenecks and Future Trends in Earthquake Prediction,” *Environ. Resour. Ecol. J.*, vol. 8, no. 1, 2024, doi: 10.23977/erej.2024.080118.
- [26] “LUTanh Activation Function to Optimize BI-LSTM in Earthquake Forecasting,” *Int. J. Intell. Eng. Syst.*, vol. 17, no. 1, hlm. 572–583, 2024, doi: 10.22266/ijies2024.0229.48.
- [27] O. N. Akarsu, “A Bibliometric Review of Earthquake and Machine Learning Research,” *Civ. Eng. Limits*, vol. 5, no. 1, hlm. 1–10, 2024, doi: 10.36937/cebel.2024.1908.
- [28] S. Tsuno, “Application of the on-Site P-Wave Earthquake Early Warning Method Based on Site-Specific Ratios of S-Waves to P-Waves to the 2016 Kumamoto Earthquake Sequence, Japan,” *Earth Planets Space*, vol. 76, no. 1, 2024, doi: 10.1186/s40623-024-01968-y.
- [29] S. Iai dan K. Ichii, “Soils and Foundations During Earthquakes,” *Soils Found.*, vol. 50, no. 6, hlm. 937–953, 2010, doi: 10.3208/sandf.50.937.
- [30] M. C. Porcu, J. C. Vielma, G. Pais, D. O. Bravo, dan J. C. V. Quintero, “Some Issues in the Seismic

- Assessment of Shear-Wall Buildings Through Code-Compliant Dynamic Analyses,” *Buildings*, vol. 12, no. 5, hlm. 694, 2022, doi: 10.3390/buildings12050694.
- [31] T. Moriya, T. Mogi, dan M. Takada, “Anomalous Pre-Seismic Transmission of VHF-band Radio Waves Resulting From Large Earthquakes, and Its Statistical Relationship to Magnitude of Impending Earthquakes,” *Geophys. J. Int.*, vol. 180, no. 2, hlm. 858–870, 2010, doi: 10.1111/j.1365-246x.2009.04461.x.
- [32] G. Cremen dan C. Galasso, “Earthquake Early Warning: Recent Advances and Perspectives,” *Earth-Sci. Rev.*, vol. 205, hlm. 103184, 2020, doi: 10.1016/j.earscirev.2020.103184.
- [33] X. Wang *dkk.*, “Small Earthquakes Can Help Predict Large Earthquakes: A Machine Learning Perspective,” *Appl. Sci.*, vol. 13, no. 11, hlm. 6424, 2023, doi: 10.3390/app13116424.
- [34] N. Guo *dkk.*, “Using Improved Support Vector Regression to Predict the Transmitted Energy Consumption Data by Distributed Wireless Sensor Network,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2020, no. 1, 2020, doi: 10.1186/s13638-020-01729-x.
- [35] G. C. Beroza, M. Segou, dan S. M. Mousavi, “Machine Learning and Earthquake Forecasting—next Steps,” *Nat. Commun.*, vol. 12, no. 1, 2021, doi: 10.1038/s41467-021-24952-6.