

PREDIKSI KEKUATAN GEMPA MENGUNAKAN MACHINE LEARNING DENGAN MODEL XGBOOST SEBAGAI LANGKAH STRATEGIS DALAM PERENCANAAN STRUKTUR BANGUNAN TAHAN GEMPA DI INDONESIA

Ari Wibowo

Teknik Arsitektur - Universitas Subang

e-mail: ariwibowo@unsub.ac.id

Abstract

The magnitude of the earthquake load acting on a building structure depends on many variables. The horizontal force, vertical force, and torsion moment that occurs due to an earthquake in a structure, are very dependent on the weight and stiffness of the structural material, configuration and structural system, the period or time of the structure's vibration, the condition of the subgrade, the earthquake area, and the behavior of the earthquake itself. So that the impact of earthquakes on building structures can be minimized, research is needed to predict the strength of earthquakes in Indonesia. Although it cannot be denied that predicting when, where and the magnitude of the next earthquake in an area or time seems quite difficult, there is an alternative for prediction using machine learning which can help predict the occurrence of an earthquake. The machine learning model used in this research is XGBoost. Meanwhile, the dataset used was taken from the BMKG catalog, where the data consists of 92,887 earthquake events in Indonesia. We separated this data into three parts as network training data, validation data and testing data with details of 80% for the training process or as many as 74309 and 20% for the testing process or as many as 18578. From the research conducted, it shows that the training stage is able to produce accuracy very well, this can be proven by looking at the model performance during training through RMSE, MAE, MSE, and R-squared losses. Where the RMSE in the training process is very small, namely 0.07 and the R-Squared (R2) value is 0.90. Meanwhile, in the data testing process, high accuracy was also obtained. This can be seen from the RMSE value of 0.01 and the R-Squared (R2) value of 0.98.

Keywords: *earthquake prediction, earthquake resistant buildings, Machine Learning, XGBoost*

Abstrak

Besarnya beban gempa yang bekerja pada struktur bangunan, tergantung dari banyak variabel. Gaya horisontal, gaya vertikal, dan momen torsi yang terjadi akibat gempa pada struktur, sangat tergantung pada berat dan kekakuan material struktur, konfigurasi, dan sistem struktur, periode atau waktu getar struktur, kondisi tanah dasar, wilayah kegempaan, serta perilaku gempa itu sendiri. Agar dampak gempa pada struktur bangunan dapat diminimalisir, maka perlu adanya penelitian dalam memprediksi kekuatan gempa di wilayah Indonesia. Walaupun tidak dapat dipungkiri memprediksi kapan, dimana dan besarnya gempa bumi berikutnya dalam wilayah atau waktu tampaknya cukup sulit, namun terdapat suatu alternative untuk prediksi menggunakan machine

learning yang dapat membantu memperkirakan terjadinya gempa. Adapun model machine learning yang digunakan dalam penelitian ini adalah XGBoost. sedangkan dataset yang digunakan diambil dari katalog BMKG, dimana data tersebut terdiri dari 92.887 kejadian gempa di Indonesia. Data tersebut kami pisahkan menjadi tiga bagian sebagai data pelatihan jaringan, data validasi dan data pengujian dengan rincian 80% untuk proses pelatihan atau sebanyak 74309 dan 20% untuk proses pengujian atau sebanyak 18578. Dari penelitian yang dilakukan, menunjukkan bahwa pada tahap pelatihan mampu menghasilkan akurasi dengan sangat baik, hal ini dapat dibuktikan dengan melihat kinerja model selama pelatihan melalui loss RMSE, MAE, MSE, dan R-squared. Dimana RMSE pada proses pelatihan sangat kecil yaitu 0.07 dan nilai R-Squared (R^2) yang bernilai 0.90. sedangkan pada proses pengujian data juga diperoleh akurasi yang tinggi. Hal ini dapat dilihat dari hasil nilai RMSE 0.01 dan nilai R-Squared (R^2) yang bernilai 0.98.

Kata Kunci: prediksi gempa, bangunan tahan gempa, Machine Learning, XGBoost

PENDAHULUAN

Gempa bumi merupakan guncangan hebat yang menjalar ke permukaan bumi yang disebabkan oleh gangguan di dalam litosfer. Gangguan ini terjadi karena di dalam lapisan kulit bumi dengan ketebalan 100 km terjadi akumulasi energy akibat pergeseran kulit bumi yang menyebabkan terlepasnya energi oleh tekanan yang dilakukan lempengan yang bergerak (Mustafa, 2010). Semakin lama tekanan akan semakin membesar dan pada akhirnya mencapai suatu keadaan dimana tekanan tidak dapat ditahan lagi sehingga terjadilah gempa bumi dan energy akan menjalar ke berbagai arah (Henny, 2009).

Gempa bumi terjadi karena diakibatkan oleh beberapa factor diantaranya gempa bumi vulkanik yang disebabkan oleh kegiatan atau aktivitas gunung berapi (Rachman, 2017), gempa runtuh atau gempa local yang terjadi apabila suatu gua di daerah topografi karst atau daerah di pertambangan runtuh atau massa batuan yang cukup besar di lereng bukit runtuh/longsor (Shaleha, et al. 2016), dan gempa buatan yang disebabkan oleh aktivitas manusia, seperti dalam kegiatan eksplorasi bahan tambang atau untuk keperluan teknik sipil dalam mencari batuan dasar sebagai material bangunan

(Bukhori, 2017).

Indonesia sendiri gempa bumi merupakan salah satu bencana alam yang sering terjadi karena Indonesia merupakan negara yang terletak di pertemuan tida lempeng tektonik utama, yaitu lempeng indo-australia, lempeng pasifik, dan lempeng Eurasia. Hal inilah yang menjadikan wilayah Indonesia dengan aktivitas seismic yang cukup tinggi (Tupan, 2020).

Gempa bumi pun merupakan salah satu bencana alam yang dapat menimbulkan kerusakan yang serius terhadap kehidupan manusia dan infrastrukturnya (Fauzi, et al. 2021). Seperti yang dilaporkan oleh badan nasional penanggulangan bencana pada gempa yang terjadi di Cianjur Jawa Barat bahwa terdapat 272 korban meninggal, 2046 orang mengalami luka-luka, total rumah rusak 56.311 dengan rincian rusak berat 22.267 unit, rusak sedang 11.836 unit dan rusak ringan 22.208 unit (BNPB, 2022). Menurut Kamurahan, S., R (2018) gempa bumi bukanlah penyebab utama terbunuhnya manusia, namun kegagalan bangunan yang banyak menimbulkan korban jiwa. Dari beberapa kasus yang berhasil didata menyebutkan rumah roboh akibat tidak mampu menahan beban horizontal selama

terjadinya gempa bumi (Tamara, 2011).

Dalam upaya mengurangi dampak buruk yang disebabkan oleh gempa bumi, penting untuk mengembangkan metode prediksi yang dapat memberikan informasi mengenai potensi terjadinya gempa bumi di wilayah Indonesia (Nur, A., M. 2010). Prediksi gempa bumi merupakan bidang yang kompleks dan sulit diprediksi secara tepat (Tantyoko, 2023). Meskipun penelitian dan teknologi terus berkembang, saat ini tidak ada metode yang dapat memberikan prediksi secara akurat dalam menentukan waktu, lokasi, dan kekuatan gempa bumi di suatu wilayah secara spesifik (Berhich, A., et al. 2022). Prediksi kapan terjadinya gempa bumi di suatu wilayah perlu dilakukan untuk mengurangi jumlah korban jiwa dan mengurangi dampak infrastrukturnya (Nur, A. 2018).

Berdasarkan uraian di atas, penulis akan melakukan penelitian mengenai prediksi gempa bumi menggunakan machine learning dengan algoritma XGBoost. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi dalam meningkatkan kemampuan prediksi gempa bumi di Indonesia, dan dapat menghasilkan model prediksi yang lebih akurat dan efektif. Dari prediksi berupa informasi besar magnitude di masa depan dapat digunakan untuk membantu pihak berwenang dan masyarakat dalam mengambil langkah-langkah mitigasi yang tepat guna mengurangi dampak buruk yang disebabkan oleh gempa bumi. Salah satunya adalah dalam perancangan struktur gedung tahan gempa di wilayah yang memiliki tingkat prediksi gempa yang tinggi.

METODE PENELITIAN

1. Alur kerja sistem

Bab ini akan membahas metodologi

penelitian yang digunakan dalam penelitian prediksi gempa bumi menggunakan algoritma XGBoost dan metode seleksi fitur. Metodologi penelitian ini mencakup langkah-langkah yang diambil dalam pengumpulan data, pengolahan data, implementasi algoritma, dan analisis hasil. Adapun penggambaran alur sistem dapat dilihat pada gambar 1.



Gambar 1. Alur kerja sistem

Adapun penjelasan mengenai alur kerja sistem di gambar 1, sebagai berikut:

- Tahap pertama adalah identifikasi masalah yang dilakukan dengan mempelajari kebutuhan dan kebermanfaatan dari penelitian yang dilakukan, menentukan permasalahan dan mengumpulkan literature penelitian dari jurnal yang berkaitan baik berupa masalah, jenis data, ataupun metode yang digunakan.
- Tahap kedua adalah studi literature yang dilakukan dengan mengkaji literature-literatur yang berkaitan seperti teori, konsep dan generalisasi terkait dengan penelitian yang dapat dijadikan sebagai landasan teori untuk penelitian ini.
- Tahap ketiga adalah pengumpulan data dengan mengumpulkan data gempa bumi di Indonesia melalui laman resmi BMKG. Yang tentu data yang diperlukan berisikan sekumpulan informasi yang berkaitan

- dengan domain gempa bumi Indonesia. Preprocessing data adalah langkah yang dilakukan untuk memuat beberapa proses yaitu pembersihan data yang diperoleh untuk mengatasi kesalahan atau nilai yang hilang, melakukan transformasi data seperti mengubah format tanggal dan waktu.
- d) Tahap keempat adalah pembagian dataset menjadi dua kelompok dimana dataset berupa data pelatihan (X_{train} dan y_{train}) dan data pengujian (X_{test} , dan y_{test}).
 - e) Tahap kelima adalah feature engineering dimana tahap ini dilakukan untuk mengekstrak fitur-fitur tambahan dari data, seperti atribut waktu dan fitur lain yang relevan dengan penelitian ini.
 - f) Tahap keenam adalah pelatihan model XGBoost menggunakan data pelatihan (X_{train} dan y_{train}) dimana model XGBoost ini digunakan untuk memprediksi besar magnitude gempa bumi berdasarkan fitur-fitur yang telah diambil dari data.
 - g) Tahap ketujuh adalah evaluasi model yang telah dilatih menggunakan metric evaluasi seperti Mean Absolute Error (MAE) dan R-squared (R^2) dengan data pengujian (X_{test} dan y_{test}). Hasil evaluasi ini akan digunakan untuk menilai sejauh mana model dapat memprediksi besar magnitude gempa dengan akurasi yang baik.
 - h) Langkah kedelapan adalah melakukan optimisasi model untuk mengoptimalkan parameter mode XGBoost dengan melibatkan penalaan hiperparameter untuk mencari parameter yang menghasilkan hasil terbaik. Dilanjutkan dengan validasi model, dengan menggunakan teknik cross-validation untuk memastikan hasil yang konsisten, dan untuk membantu mengetahui dan

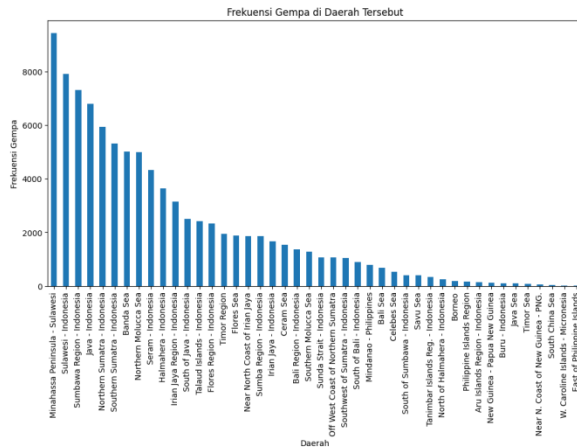
mengidentifikasi apakah model cenderung overfitting atau underfitting.

- i) Setelah menggunakan model XGBoost yang telah dilatih untuk memprediksi besar magnitude gempa bumi pada data baru dengan fitur yang sama seperti tanggal, waktu, koordinat geografis, dan kedalaman.

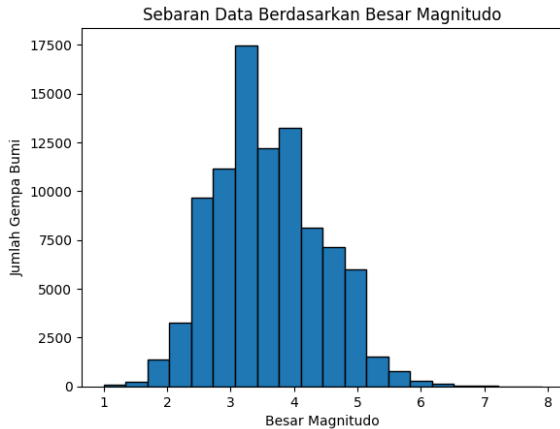
2. Data set

Dataset adalah kumpulan data yang terorganisir dalam format yang terstruktur. Dataset ini berisikan atas kumpulan informasi yang berkaitan dengan domain gempa bumi di Indonesia. Dataset digunakan dalam berbagai bidang seperti ilmu computer, statistic, ilmu data, dan pembelajaran mesin. Pada penelitian ini, dataset digunakan untuk melatih dan menguji model atau algoritma pembelajaran mesin. Sedangkan secara keseluruhan dataset ini diambil dari katalog resmi BMKG dari tanggal 1 november 2008 hingga 15 April 2023.

Data tersebut terdiri dari 92887 kejadian gempa yang terjadi di Indonesia dan berisikan beberapa variabel, diantaranya: tanggal terjadinya gempa bumi, ot atau waktu yang spesifik dalam jam, menit, detik, lokasi atau letak koordinat dalam latitude dan longitude yang berkisar antara 6° LU hingga 11° LS sedangkan bujurnya berkisar antara 142° BT hingga 94° BT, kedalaman terjadinya gempa dalam km, dan besar magnitude setiap gempa bumi terjadi yang berkisar antara 1 hingga 9,5, untuk sebaran data berdasarkan frekuensi gempa di Indonesia selengkapnya dapat dilihat pada gambar 2, sebaran data berdasarkan besar magnitude dapat dilihat pada gambar 3, sebaran data lokasi terjadinya gempa di Indonesia dapat dilihat pada gambar 4, dan sebagian contoh data dari keseluruhan data gempa bumi di Indonesia dapat dilihat pada tabel 1.



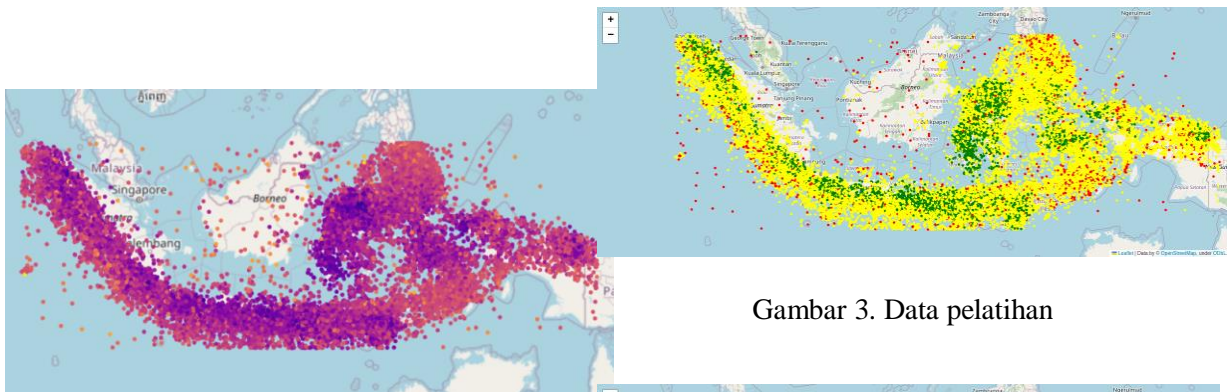
Gambar 2. Sebaran data berdasarkan frekuensi gempa di Indonesia



Gambar 3. Sebaran magnitude gempa di Indonesia

Tabel 1. Contoh sebagian data gempa bumi di Indonesia

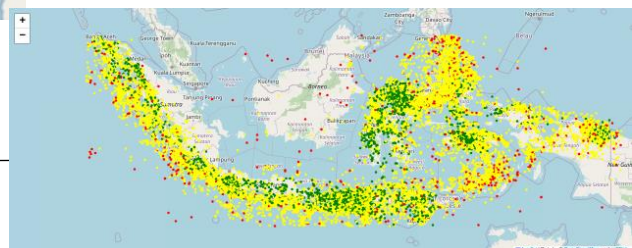
Dari keseluruhan data yang diperoleh ini, nantinya akan digunakan untuk pelatihan model XGBoost dan pengujian dari model XGBosst. Adapun tahapn ini diistilahkan sebagai split dataset dan dengan perbandingan untuk jumlah masing-masing antara data train dan data tes ini terhadap keseluruhan dataset yang digunakan yaitu sebanyak 92.887 titik gempa di Indonesia adalah 80% digunakan untuk pelatihan sehingga berjumlah 74.309 dan 20% digunakan untuk pengujian sehingga jumlahnya sebanyak 18.578 data. Adapun untuk lebih jelasnya mengenai sebaran gempa bumi di Indonesia baik untuk data pelatihan maupun data pengujian, dapat dilihat pada gambar berikut ini.



Gambar 3. Data pelatihan

Gambar 3. Sebaran data lokasi gempa bumi

	tgl	ot	lat	lon	depth	mag	remark	st
0	2008/11/01	21:02:43.058	-9.18	119.06	10	4.9	Sumba Region - Indonesia	
1	2008/11/01	20:58:50.248	-6.55	129.64	10	4.6	Banda Sea	
2	2008/11/01	17:43:12.941	-7.01	106.63	121	3.7	Java - Indonesia	
3	2008/11/01	16:24:14.755	-3.30	127.85	10	3.2	Seram - Indonesia	
4	2008/11/01	16:20:37.327	-6.41	129.54	70	4.3	Banda Sea	



Gambar 4. Data pengujian

3. Prediksi

Setelah melalui proses seperti yang disebutkan di bab sebelumnya, dengan melalui proses terakhir yaitu pelatihan sudah dilakukan, berjalan baik, dan akurasi sudah cukup baik dari penggunaan model XGBoost. Dari model yang sudah terlatih untuk mengenali pola dalam data pelatihan yang sesuai dengan magnitude gempa, maka langkah berikutnya adalah melakukan proses prediksi besar magnitude gempa bumi pada data baru dengan fitur yang sama, seperti fitur tanggal, waktu, koordinat, geografis, dan kedalaman. Prediksi ini akan memperkirakan magnitude gempa di lokasi dan waktu yang telah ditentukan dalam data inputan.

HASIL DAN PEMBAHASAN

1. Implementasi model XGBoost

Pada dasarnya dalam penelitian deteksi gempa bumi di Indonesia, dengan menggunakan XGBoost terdapat beberapa tahapan penting yaitu penggunaan dataset, pembersihan data, pembagian data menjadi data pelatihan dan data pengujian, pelatihan model, pengujian model, validasi dan melakukan prediksi. Dari proses validasi ini nanti akan digunakan sebagai tolak ukur terhadap proses berikutnya yaitu proses pengujian maupun proses prediksi. Apabila memberikan hasil yang baik, maka jaringan tersebut dapat digunakan untuk melakukan klasifikasi data dengan data tes.

2. Proses pelatihan dan pembentukan model

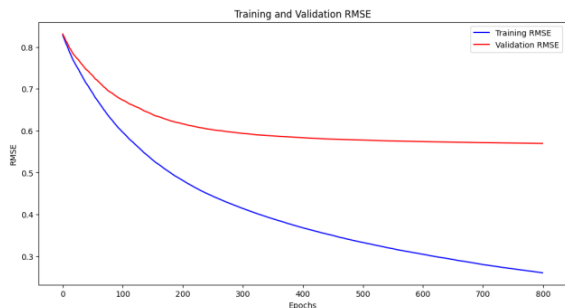
Data latih yang digunakan adalah 80% dari total keseluruhan data yang sudah terkumpul. Sehingga didapat data latih sebanyak 74309 dengan fitur untuk pemilihan kolom inputan dan kolom target. Dimana kolom input berupa tahun, bulan, hari, jam, menit, detik, garis lintang, garis bujur, dan kedalaman, sedangkan untuk kolom target adalah magnitude gempa bumi. Komputasi dilakukan dengan menggunakan mode single GPU. Kemudian untuk proses training menggunakan hyperparameter untuk model XGBoost sebagai berikut:

```
n_estimators : 50-400/50
max_depth    : 3-10
learning_rate : 0.001, 0.01,
              0.1, 0.2, 0.3
sub_sampel   : 0.6-1.0
colsample_bytree : 0.6- 1.0
gamma        : 0, 0.1, 0.2,
              0.3, 0.4
n_iterasi    : 50
```

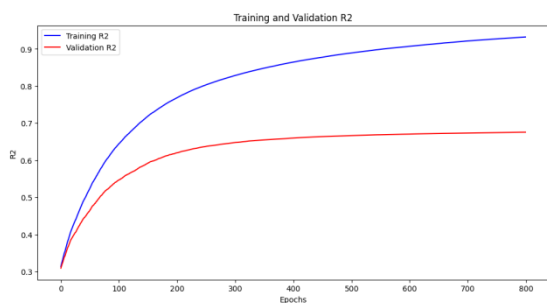
Selengkapnya dapat dilihat pada gambar 3. Dari gambar 3 ini dapat disimpulkan bahwa struktur XGBoost telah sesuai dengan parameter yang dijelaskan di atas. Sedangkan untuk jejak proses trainingnya dapat dilihat pada grafik dari akurasi dan kesalahan selama proses training berlangsung baik dari matriks evaluasi atau RMSE (Root Mean Square Error) maupun koefisien determinasi atau R^2 (Coefficient of determination), seperti pada gambar 4 dan gambar 5.

```
XGBRegressor
XGBRegressor(base_score=None, booster=None, callbacks=None,
              colsample_bylevel=None, colsample_bynode=None,
              colsample_bytree=None, device=None, early_stopping_rounds=None,
              enable_categorical=False, eval_metric=None, feature_types=None,
              gamma=None, grow_policy=None, importance_type=None,
              interaction_constraints=None, learning_rate=None, max_bin=None,
              max_cat_threshold=None, max_cat_to_onehot=None,
              max_delta_step=None, max_depth=None, max_leaves=None,
              min_child_weight=None, missing=nan, monotone_constraints=None,
              multi_strategy=None, n_estimators=None, n_jobs=None,
              num_parallel_tree=None, random_state=None, ...)
```


Gambar 3. Struktur model XGBoost



Gambar 4. Grafik RMSE dari proses training dan validasi



Gambar 5. Grafik R² dari proses training dan validasi

Dari grafik RMSE dan R² terdapat informasi mengenai keakurasian dari model yang digunakan untuk pelatihan maupun tahap validasi. Data pelatihan sendiri digunakan untuk melatih model machine learning yang digunakan, dimana data ini digunakan untuk model dalam memahami hubungan antara fitur (input) dan label (output). Sedangkan data validasi digunakan selama proses pelatihan untuk memantau sejauh mana model dapat menggeneralisasi dari data pelatihan ke data yang tidak pernah dilihat sebelumnya. Tahap ini memungkinkan untuk membantu dalam penyesuaian hyperparameter dan pencegahan overfitting, dan digunakan untuk mengukur kinerja model pada data yang tidak digunakan dalam pelatihan.

Oleh karena itu, untuk mengevaluasi proses

pelatihan maupun proses validasi, salah satu cara yang dapat digunakan adalah dengan melihat kinerja melalui grafik RMSE dan R-squared (R²). Semakin rendah nilai RMSE, maka akan semakin baik model dalam memprediksi data. Sebaliknya, semakin tinggi RMSE, semakin besar kesalahan prediksi model Machine learning yang digunakan. Berbeda dengan R², dimana nilai R² berkisar antara 0 hingga 1, nilai 0 berarti model sama sekali tidak dapat menjelaskan variasi dalam data, sedangkan nilai 1 berarti model dapat menjelaskan semua variasi dalam data. Sedangkan MSE (Mean Squared error) adalah metrik yang dipakai untuk mengukur rata-rata dari kesalahan kuadrat selama proses pelatihan model, dan MSE ini digunakan untuk menghitung sejauh mana model mendekati data pelatihan. Semakin rendah nilai MSE, menunjukkan model memiliki sedikit kesalahan kuadrat dalam memprediksi data pengujian. Sedangkan untuk MAE (Mean Absolute Error) adalah metrik yang mengukur rata-rata kesalahan absolut (tanpa kuadrat) dalam mengevaluasi sejauh mana model berkinerja pada data pelatihan. Semakin kecil nilai MAE maka dapat menunjukkan bahwa model memiliki kesalahan absolute yang rendah dalam memprediksi data pengujian.

Jika dilihat dari grafik yang diperoleh baik RMSE maupun R² dapat disimpulkan bahwa pelatihan model terhadap data training maupun validasi memberikan akurasi yang sangat baik. Hal ini dapat dibuktikan bahwa grafik RMSE menunjukkan adanya penurunan secara signifikan hingga iterasi ke-800. Baik Training RMSE yang bernilai 0.07 dan validasi RMSE yang bernilai 0.63. selain itu diperkuat oleh grafik R² yang menyatakan bahwa proses training dan validasi mengalami kenaikan hingga mendekati 1 hingga itersi ke-800. Dimana Training R² sebesar 0.90 dan

validasi R^2 sebesar 0.53. Sedangkan nilai MAE pada tahap pelatihan adalah 0.19 sedangkan validasi adalah 0.45, untuk nilai MSE pada proses pelatihan adalah 0.07 dan untuk validasi adalah 0.32. Atau lebih jelasnya dapat dilihat pada tabel 2 berikut ini.

Evaluasi model regresi	Training	Validasi
RMSE	0.07	0.63
R^2	0.90	0.53
MAE	0.19	0.45
MSE	0.07	0.32

Tabel 2. Evaluasi model XGBoost pada proses pelatihan

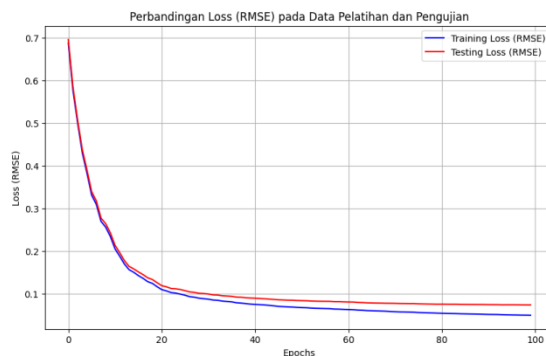
3. Proses pengujian dan proses prediksi

Proses pengujian dan proses prediksi merupakan dua tahap penting dalam evaluasi dan penggunaan model machine learning. Perbedaan utama antara keduanya adalah bahwa proses pengujian adalah tahap evaluasi menggunakan data uji yang sudah ditentukan, sedangkan proses prediksi adalah tahap dimana model sebenarnya digunakan untuk membuat prediksi atas data yang belum dilihat. Dan proses prediksi sendiri merupakan hasil akhir dari proses pelatihan dan evaluasi model, yang bertujuan untuk memberikan nilai prediksi yang berguna dalam konteks aplikasi dunia nyata. Selama proses pengujian, metrik evaluasi seperti Mean Squared Error (MSE), Mean Absolute Error (MAE), R-squared (R^2), akurasi, presisi, recall, atau F1-score dapat digunakan untuk mengukur kinerja model sesuai dengan jenis masalah yang dihadapi.

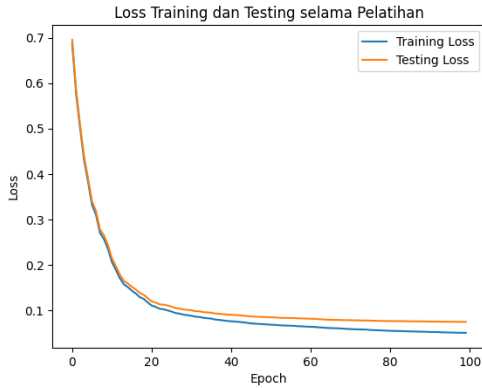
Skenario seperti di atas, yang digunakan pada proses pengujian dan proses pendektaksian pada penelitian yang dilakukan, yaitu

mengenai gempa bumi di Indonesia. Proses pengujian dilakukan dengan menggunakan data uji dengan presentase 20% terhadap total dataset yang telah dikumpulkan, atau jika data totalnya sebanyak 92.887 data, maka data yang digunakan untuk proses pengujian sebanyak 18.557 data. Dan hasil pengujian didapatkan nilai evaluasi untuk MSE (Mean Square Error) yang sangat rendah yaitu 0.01, dan nilai R^2 sangat tinggi yaitu 0.98 dan hampir mendekati 1.

Hal ini dapat terlihat pada beberapa grafik perbandingan nilai Loss RMSE pada data pelatihan dan data pengujian, seperti terlihat pada gambar 6 berikut ini.

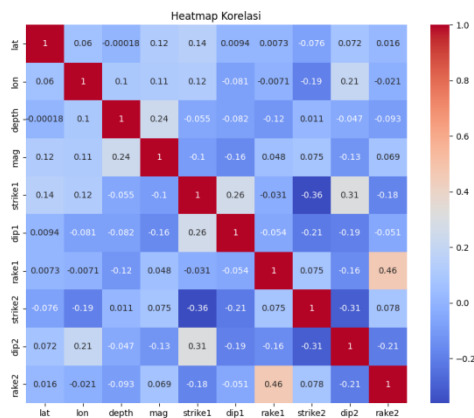


Gambar 6. Grafik perbandingan Loss RMSE dari proses pelatihan dan pengujian



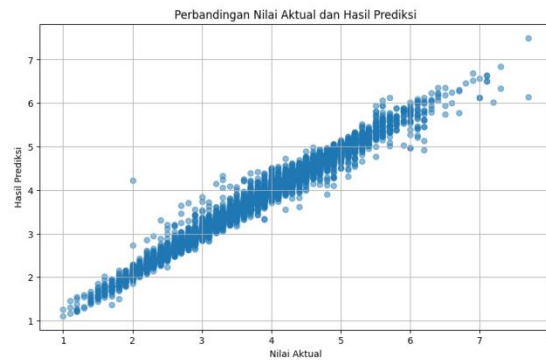
Gambar 7. Grafik loss dari proses training dan testing

Dari grafik gambar 6 dan 7 merupakan bentuk visualisasi bagaimana loss berubah selama pelatihan. Dari gambar tersebut dapat disimpulkan bahwa loss atau kesalahan dalam memprediksi data pelatihan dan pengujian semakin menurun secara signifikan dan pada mencapai titik terendah di iterasi ke-100. dengan nilai loss dibawah 0.1. sehingga dapat dipastikan bahwa model yang digunakan dan pendekatan yang dilakukan secara keseluruhan mendekati nilai sebenarnya atau ground truth serta dapat memahami data dan melakukan prediksi yang akurat. Sedangkan untuk memahami hubungan atau asosiasi antara variabel dalam dataset serta mengetahui pengaruh kuat satu variabel dengan variabel lain, dapat dilihat pada gambar 8 di bawah ini.



Gambar 8. Visualisasi heatmap korelasi tiap variabel

Dari gambar visualisasi heatmap korelasi dapat disimpulkan bahwa atribut yang memiliki warna biru tua adalah atribut yang kurang relevan dengan atribut outputnya. Sedangkan untuk atribut warna merah tua adalah atribut yang relevan dengan atribut outputnya. Dan jika diamati pada gambar tersebut, warna merah ditiap atribut inputan tepat pada tiap atribut outputan sehingga dapat dipastikan bahwa setiap atribut inputan dan outputannya menunjukkan korelasi yang sangat kuat atau memiliki hubungan linier sempurna antara atribut inputan dengan outputan. Sehingga dapat berdampak signifikan terhadap variabel target. Kemudian untuk mengetahui perbandingan antara nilai aktual dengan hasil prediksi dari model, dapat dilihat pada grafik scatter plot seperti pada gambar 9 di bawah ini.



Gambar 9. Scatter plot antara nilai aktual dan hasil prediksi

Dari grafik scatter plot pada gambar 9 dapat disimpulkan sebaran titik-titik dengan warna biru tua merupakan nilai actual, dan biru muda adalah nilai prediksi. Terlihat untuk warna biru muda mendekati atau cocok dengan warna biru tua. Selain terlihat mendektai atau cocok, pada gambar tersebut, dapat dilihat bahwa untuk kedua warna juga mendekati cenderung semakin mendekati garis diagonal maka dapat dipastikan semakin baik kinerja model yang digunakan untuk memprediksi gempa bumi di Indonesia.

KESIMPULAN

Seperti pada penjelasan di bab sebelumnya bahwa penelitian yang dilakukan yaitu prediksi gempa bumi di Indonesia dengan menggunakan machine learning dengan model XGBoost telah berhasil dilakukan dan menghasilkan akurasi yang sangat tinggi baik pada tahap pelatihan maupun tahap pengujian serta berdasarkan atribut-atribut yang digunakan.

Hal ini dapat dibuktikan bahwa nilai RMSE pada tahap pelatihan memiliki nilai yang sangat rendah, yaitu 0.07 dan diperkuat oleh grafik R^2 yang menyatakan bahwa proses training mengalami kenaikan hingga mendekati 1 hingga itersi ke-800. Dimana Training R^2 sebesar 0.90. diperkuta dengan nilai MAE pada tahap pelatihan adalah 0.19 untuk nilai MSE pada proses pelatihan adalah 0.07. Sedangkan pada tahap pengujian nilai MSE yang rendah menunjukkan bahwa model yang dikembangkan memiliki kesalahan yang sangat kecil dalam memprediksi data uji. Dan dapat dikatakan model hampir sempurna dalam memodelkan data. Sedangkan nilai R^2 yang mendekati 1 menandakan bahwa model dapat menjelaskan sekitar 98% dari variasi dalam data pengujian. Model ini sangat cocok dengan data dan sangat baik dalam menggambarkan pola data. Dengan kata lain, model ini dianggap sangat efektif dalam memodelkan data gempa bumi dan memprediksi magnitude dengan sangat baik.

Dengan diketahuinya prediksi besaran magnitude dan gaya gempa yang terjadi di Indonesia, merupakan suatu usaha dalam merancang struktur bangunan tahan gempa. Dimana besarnya gaya gempa yang mengenai pada masing-masing tingkat dari bangunan gedung tergantung dari berat dan ketinggian

tingkat suatu gedung. Karena berat dan ketinggian tingkat merupakan beban gempa yang bekerja pada struktur bangunan dan disebut sebagai gaya inersia.

Sehingga dalam mempertimbangkan bangunan tahan gempa, harus mempertimbangkan beberapa factor, diantaranya berat atau massa bangunan dan percepatan gempa, dimana percepatan gempa ini merupakan factor utama ketika akan mempertimbangkan bangunan yang tahan terhadap gempa. Hal-hal lain yang perlu diperhatikan selain ketiga factor utama tersebut antara lain; tata letak dari struktur bangunan, perencanaan kapasitas dengan konsep *strong column - weak beam*, pendetailan yang baik dari elemen-elemen struktur, memperhitungkan cara mendistribusikan massa dari bangunan ketika terjadi gempa, kekakuan dari sistem struktur bangunan, dan perilaku dari getaran gempa. Walaupun pada dasarnya perilaku dari getaran gempa paling sulit ditentukan secara tepat karena sifatnya yang acak (random). Akan tetapi jika bangunan memenuhi persyaratan-persyaratan di atas, maka diharapkan perencanaan struktur di daerah rawan gempa dapat mengurangi dampak ketika terjadi gempa bumi di Indonesia.

SARAN

Dalam proses penelitian yang dilakukan yaitu prediksi gempa bumi di Indonesia dengan menggunakan machine learning dengan model XGBoost ditemukan beberapa kendala. Diantaranya karena pada saat proses penelitian menggunakan google colabs versi free, sehingga mengalami keterbatasan sumber daya termasuk CPU, RAM, dan GPU, menyebabkan runtime error, sehingga disarankan beralih menggunakan akun google colabs yang berbayar, optimalisasi kode akun google colabs dengan lebih efektif,

menyimpan secara berkala dan mengamankan data serta proses yang telah dijalankan, mengatasi kesalahan atau bug pada kode, memahami batasan waktu dan sumber daya yang tersedia di colabs, dan jika memungkinkan untuk menggunakan atau menjalankan proyek di PC local.

Selain itu, sebagai bentuk evaluasi untuk penelitian selanjutnya, untuk meningkatkan akurasi dan menurunkan nilai loss dengan melakukan beberapa hal seperti menambahkan data citra, regularisasi atau pengendalian kompleksitas model, dropout atau pengurangan overfitting, menambah lapisan konvolusi dan menambahkan fungsi optimasi. Dan dapat juga menggunakan host platform lainnya.

DAFTAR PUSTAKA

- Berhich, A. et al. (2022). A location-dependent earthquake prediction using recurrent neural network algorithms. *Soil Dynamics and Earthquake Engineering*, Vol. 161, DOI: <https://doi.org/10.1016/j.soildyn.2022.107389>
- BNPB. 2022. [UPDATE] 268 Warga Meninggal Dunia Akibat Gempa Cianjur. <https://bnpb.go.id/berita/-update-268-warga-meninggal-dunia-akibat-gempa-cianjur->. Diakses tanggal 12 Oktober 2023.
- Bukhori. 2017. Pengertian Gempa Bumi. <https://karyapemuda.com/pengertian-gempa-bumi/>. Diakses tanggal 24 Oktober 2023.
- Fauzi, M. et al. (2021). Dampak Bencana Gempabumi dan Tsunami di Kawasan Pesisir Lere Kota Palu. *Jurnal Pembangunan Wilayah dan Kota*, Vol. 17, No. 1, 2021, pp 16-24, DOI: : <https://doi.org/10.14710/pwk.v17i1.29967>.
- Kamurahan, S., R., et al. (2018). Struktur dan Konstruksi Rumah Panggung Masyarakat Kampung Jawa Tondano (Jaton) Ditinjau dari Prinsip-Prinsip Bangunan Tahan Gempa. *Media Matrasain*, Vol. 15, No. 1 (2018), DOI: <https://doi.org/10.35792/matrasain.v15i1.21183>.
- Mustafa, B. (2010). Analisis Gempa Nias dan Gempa Sumatera Barat dan Kesamaannya Yang Tidak Menimbulkan Tsunami. *Jurnal Ilmu Fisika*. Vol. 2, No. 1, 44–50 (2010).
- Noor, A. (2018). Perbandingan Algoritma Support Vector Machine Biasa dan Support Vector Machine Berbasis Particle Swarm Optimization untuk Prediksi Gempa. 2018.
- Nur, A.,M., (2010). Gempa Bumi, Tsunami dan Mitigasinya. Balai Informasi dan Konservasi Kebumian Karangsembung – LIPI, Kebumen. Vol. 7, No. 1 (2010).
- Henny, P. A., et al. (2009). Studi Tentang Kerusakan Infrastruktur Keairan Akibat Gempa Tektonik di Kabupaten Klaten. Semarang: Universitas Islam Sultan Agung.
- Rachman, A., et al. (2017). Penentuan Magnitudo Gempa Bumi dengan Menganalisis Amplitudo Anomali Manetik Prekursor Gempa Bumi Dan Jarak Hypocenter. *Jurnal Teknik ITS* Vol 6, No. 2 (2017).
- Shaleha, A., et al. (2016). Identifikasi Struktur Lapisan Tanah Daerah Rawan Longsor di Kecamatan Banyubiru Kabupaten Semarang dengan Metode Horizontal To Vertical Spectral Ratio (HVSR). *Unnes Physics Journal*, Vol. 5 No. 2.
- Tamara, M. (2011). Evaluasi Kerusakan

Bangunan Akibat Gempa Besar.
Jurnal Ilmiah MEDIA
ENGINEERING, Vol. 1, No. 1
(2011), pp 1-9.

Tantyoko, H. et al. (2023). Prediksi Potensial
Gempa Bumi Indonesia Menggunakan
Metode Random Forest dan Feature
Selection. *Idealis: Indonesia Journal
Information System*, Vol. 6, NO. 2
(2023).

DOI:<https://doi.org/10.36080/idealis.v6i2.3036>.

Tupan, et al. (2020). Analisis Bibliometrik
Publikasi Ilmiah Tentang Prediksi
Gempa Bumi Berbasis Data Scopus
Periode 2015-2020. *Libraria Jurnal
Perpustakaan*, vol. 8, no. 1, 2020.

DOI:

<http://dx.doi.org/10.21043/libraria.v8i1.7183>.