



Optimasi lstm dengan nesterov adaptive moment estimation optimizer untuk prediksi kekuatan magnitudo gempa bumi di pulau sumatera

Aditya Darma¹

Email: ¹210401174@student.umri.ac.id

¹Teknik Informatika, Ilmu Komputer, Universitas Muhammadiyah Riau

Diterima: - | Direvisi: - | Disetujui: -

©2025 Program Studi Teknik Informatika Fakultas Ilmu Komputer,
Universitas Muhammadiyah Riau, Indonesia

Abstrak

Penelitian ini bertujuan untuk mengoptimalkan model LSTM dengan optimasi NADAM untuk memprediksi kekuatan magnitudo gempa bumi. *Dataset* yang digunakan terdiri dari 13.419 data gempa bumi yang mencakup wilayah sumatera bagian barat, utara, dan selatan. Tahapan penelitian meliputi penyaringan data, cleaning data, pembagian data (*data splitting*), normalisasi dengan *Robust Scaler*, hingga pemodelan menggunakan LSTM. Data dibagi menjadi tiga skenario pembagian data latih dan data uji, yaitu 90:10, 80:20, dan 70:30, dengan variasi timestep 500 dan 1000. Proses optimasi model melibatkan penyesuaian *hyperparameter* seperti jumlah unit LSTM, *learning rate*, *batch size*, dan jumlah *epoch*. Hasil eksperimen menunjukkan bahwa kombinasi data latih 80%, data uji 20%, timestep 1000, batch size 32, dan epoch 60 memberikan hasil terbaik dengan nilai MAE sebesar 0,5269, MAPE 15,81%, MSE 0,4586, dan RMSE 0,6772. Penelitian ini menunjukkan bahwa algoritma LSTM dengan optimasi NADAM memiliki kinerja yang baik dalam memprediksi kekuatan gempa bumi, serta dapat digunakan untuk mendukung mitigasi bencana di masa depan.

Kata kunci: *LSTM, NADAM, robust scaler, prediksi, magnitudo gempa*

Lstm optimization with nesterov adaptive moment estimation optimizer for prediction of earthquake magnitude strength on the island of Sumatra

Abstract

This study aims to optimize the LSTM model with NADAM optimization to predict earthquake magnitude strength. The dataset used consists of 13,419 earthquake data covering the western, northern, and southern Sumatra regions. The research stages include data filtering, data cleaning, data splitting, normalization with Robust Scaler, and modeling using LSTM. The data was divided into three scenarios of training and test data division, namely 90:10, 80:20, and 70:30, with timestep variations of 500 and 1000. The model optimization process involves adjusting hyperparameters such as the number of LSTM units, learning rate, batch size, and number of epochs. The experimental results show that the combination of 80% training data, 20% test data, 1000 timesteps, batch size 32, and 60 epochs provides the best results with MAE values of 0.5269, MAPE 15.81%, MSE 0.4586, and RMSE 0.6772. This research shows that the LSTM algorithm with NADAM optimization has good performance in predicting earthquake strength, and can be used to support disaster mitigation in the future.

Keywords: *LSTM, NADAM, robust scaler, prediction, earthquake magnitude*

1. PENDAHULUAN

Latar Belakang

Indonesia merupakan negara yang terletak pada pertemuan tiga lempeng tektonik utama yang terdiri dari lempeng Indo-Australia, Lempeng Pasifik, dan Lempeng Eurasia Sehingga menjadikan Indonesia sebagai negara dengan wilayah yang memiliki aktivitas seismik tinggi dan rawan terjadi bencana gempa bumi akibat kompleksitas tektonik yang tinggi tersebut [1]. Gempa bumi adalah fenomena yang terjadi akibat pergerakan lempeng tektonik yang dapat memiliki dampak merusak. Gempa bumi merupakan salah satu bencana alam yang sulit untuk dihindari dan dapat menyebabkan kerusakan fisik yang besar, dampak di bidang ekonomi yang signifikan serta dapat menghilangkan nyawa manusia. Salah satu wilayah di Indonesia yang rawan terkena dampak dari bencana alam ini adalah pulau Sumatera. Sumatera merupakan sebuah pulau yang dikategorikan sebagai wilayah tektonik aktif dunia. Menurut kementerian ESDM, 6 dari 25 wilayah rawan gempa di Indonesia terletak di pulau Sumatera yang diantaranya adalah Aceh, Jambi, Bengkulu, Lampung, Sumatera Barat, dan Sumatera Utara. Selain itu, dalam kurun waktu 20 tahun terakhir terlihat bahwa gempa dengan skala kekuatan yang besar sering terjadi di pulau Sumatera, hal ini diakibatkan karena pengaruh dari letak geografis wilayah ini yang dilalui oleh sesar aktif, zona subduksi, dan jalur gunung berapi.

Badan Meteorologi Klimatologi dan Geofisika (BMKG) merupakan lembaga pemerintah Non Departemen (LPND) yang bertugas melaksanakan tugas pemerintahan di bidang Meteorologi, Klimatologi, Kualitas Udara dan Geofisika sesuai dengan ketentuan perundang-undangan yang berlaku termasuk didalamnya memonitoring aktivitas gempa bumi yang terjadi di wilayah Indonesia. Menurut BMKG berdasarkan data yang dikumpulkan selama periode 2009–2018 menunjukkan bahwa Indonesia telah diguncang gempa sebanyak 5937 kali dan 36% diantaranya adalah gempa dengan kekuatan besar. Beberapa bencana gempa bumi dengan skala kekuatan besar yang terjadi di pulau Sumatera diantaranya adalah gempa Nias pada tahun 2005 dengan kekuatan 8.6 M dan gempa Aceh pada tahun 2004 yang menimbulkan bencana Tsunami Sehingga mengakibatkan banyak korban jiwa sekitar 250.000 orang serta kerusakan infrastruktur dan memicu terjadinya gempa dan tsunami di beberapa negara tetangga seperti Sri Lanka, Thailand, dan India [2].

Salah satu upaya penting yang dapat dilakukan dalam mengurangi dampak buruk yang disebabkan oleh gempa bumi adalah dengan memanfaatkan kecerdasan buatan untuk mengetahui kekuatan gempa yang mungkin terjadi melalui prediksi. Prediksi atau peramalan adalah suatu kegiatan yang dilakukan seorang peneliti untuk memperkirakan kejadian yang terjadi di masa depan dengan menggunakan metode atau pendekatan ilmiah [3]. Tingkat keakuratan dari suatu prediksi yang dilakukan ditentukan oleh seberapa jauh perbedaan antara data asli dengan data yang dilakukan prediksi. Pada pembuatan model prediksi kekuatan gempa, data yang akan digunakan berupa *datetime*, *latitude*, *longitude*, *depth*, dan *magnitudo*. Kemudian, penulis menggunakan algoritma *Long Short Term Memory* (LSTM) dengan optimasi *Nesterov Adaptive Moment Estimation* (NADAM) untuk mengoptimalkan kinerja model prediksi dalam memprediksi kekuatan gempa yang akan terjadi serta menggunakan beberapa algoritma untuk mengevaluasi tingkat akurasi model. LSTM dipilih karena merupakan salah satu arsitektur *Recurrent Neural Network* (RNN) yang cocok digunakan untuk data berbentuk *time series* dan memiliki kemampuan dalam menyimpan pola data untuk melakukan prediksi serta memiliki sensitifitas untuk melakukan seleksi dan mempelajari mengenai pola data Sehingga bisa menyimpulkan data mana yang akan di eliminasi dan data mana yang tetap dipertahankan. beberapa penelitian terdahulu mengenai perbandingan hasil prediksi LSTM dengan algoritma lain seperti *Extreme Gradient Boosting* (XGBoost) dalam memprediksi harga pembukaan saham dan perbandingan LSTM dan *Autoregressive Interated Moving AVERAGE* (ARIMA) untuk memprediksi kebutuhan bahan baku menunjukkan bahwa LSTM unggul dalam hasil prediksi dibandingkan

kedua algoritma tersebut. *Nesterov Adam* (NADAM) sendiri merupakan pengembangan dari algoritma *Adaptive Moment Estimation* (ADAM) yang bekerja dengan memperbarui bobot untuk menghaluskan gradien yang berdampak dalam mempercepat proses training serta meningkatkan akurasi.

Oleh karena itu, Penggunaan algoritma LSTM dengan optimasi NADAM akan menjadi pilihan yang menarik untuk digunakan dalam memprediksi kekuatan gempa di pulau Sumatera. Hasil dari model prediksi kekuatan gempa ini merupakan salah satu solusi untuk meminimalisir kemungkinan dampak buruk dari bencana alam yang terjadi. Dengan diketahuinya kekuatan gempa yang mungkin akan terjadi di masa depan di pulau Sumatera, berbagai pihak yang berkaitan dengan kebencanaan pada tiap wilayah di pulau ini dapat mengambil respons cepat untuk meningkatkan kewaspadaan yang dapat berupa perencanaan bangunan, kota serta dapat mengedukasi masyarakat di wilayah tersebut.

Tinjauan literatur singkat

Pada pembuatan model prediksi kekuatan gempa, data yang akan digunakan berupa *datetime*, *latitude*, *longitude*, *depth*, dan *magnitudo*. Kemudian, penulis menggunakan algoritma *Long Short Term Memory* (LSTM) dengan optimasi *Nesterov Adaptive Moment Estimation* (NADAM) untuk mengoptimalkan kinerja model prediksi dalam memprediksi kekuatan gempa yang akan terjadi serta menggunakan beberapa algoritma untuk mengevaluasi tingkat akurasi model. LSTM dipilih karena merupakan salah satu arsitektur *Recurrent Neural Network* (RNN) yang cocok digunakan untuk data berbentuk *time series* dan memiliki kemampuan dalam menyimpan pola data untuk melakukan prediksi serta memiliki sensitifitas untuk melakukan seleksi dan mempelajari mengenai pola data. Sehingga bisa menyimpulkan data mana yang akan di eliminasi dan data mana yang tetap dipertahankan. Beberapa penelitian terdahulu mengenai perbandingan hasil prediksi LSTM dengan algoritma lain seperti *Extreme Gradient Boosting* (XGBoost) dalam memprediksi harga pembukaan saham dan perbandingan LSTM dan *Autoregressive Interated Moving AVERAGE* (ARIMA) untuk memprediksi kebutuhan bahan baku menunjukkan bahwa LSTM unggul dalam hasil prediksi dibandingkan kedua algoritma tersebut. *Nesterov Adam* (NADAM) sendiri merupakan pengembangan dari algoritma *Adaptive Moment Estimation* (ADAM) yang bekerja dengan memperbarui bobot untuk menghaluskan gradien yang berdampak dalam mempercepat proses training serta meningkatkan akurasi.

Oleh karena itu, Penggunaan algoritma LSTM dengan optimasi NADAM akan menjadi pilihan yang menarik untuk digunakan dalam memprediksi kekuatan gempa di pulau Sumatera

Alasan diadakan penelitian ini

Beberapa alasan yang memperkuat diadakannya penelitian ini antara lain adalah sebagai berikut :

1. Pulau Sumatera merupakan wilayah rawan terkena gempa bumi yang dapat memberikan dampak kerusakan bahkan korban jiwa sehingga menjadi tantangan serius dalam upaya mitigasi dan pemulihan.
2. Kurangnya model prediksi berbasis data yang mampu menyediakan informasi akurat terkait potensi kekuatan gempa bumi yang akan terjadi yang sangat dibutuhkan untuk mempercepat respons bencana dan mengurangi dampak kerugian yang mungkin terjadi

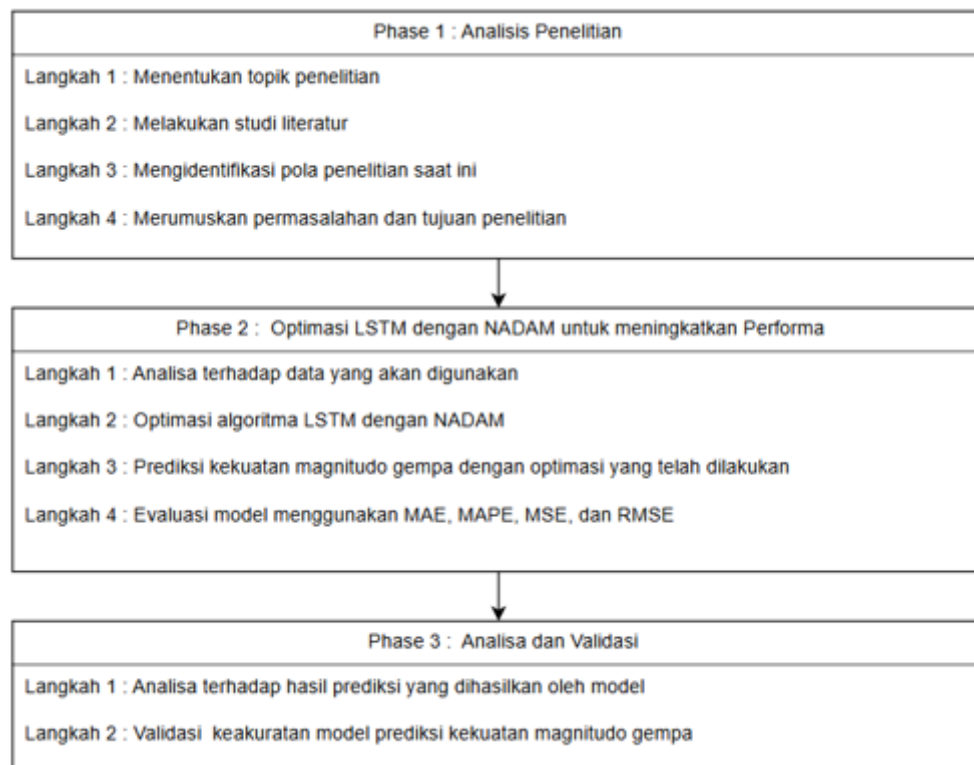
Pertanyaan Tujuan

Pertanyaan tujuan dari penelitian ini adalah Bagaimana cara mengoptimalkan algoritma Long Short-Term Memory (LSTM) dengan menggunakan optimasi Nesterov Adaptive Moment Estimation (NADAM) untuk meningkatkan akurasi prediksi kekuatan magnitudo gempa bumi di Pulau Sumatera?

2. METODE PENELITIAN

2.1 Tahapan Metode Penelitian

Metodologi penelitian merupakan sebuah tahapan yang disusun dengan sistematis dan logis dalam menjalankan suatu penelitian guna untuk mencapai tujuan dari penelitian tersebut. Tahapan yang dilakukan adalah menentukan topik penelitian, studi literatur, pengumpulan data, preprocessing data, *data splitting*, *feature engineering*, perancangan model prediksi dengan LSTM, pelatihan dan validasi model, evaluasi model dengan MAE, MAPE, MSE dan RMSE kemudian visualisasi hasil prediksi. Tahapan penelitian ini dapat dilihat pada kerangka kerja pada Gambar 1 berikut.



Gambar 1. Tahapan Metodologi Penelitian

2.2 Analisis Penelitian

Pada tahap ini langkah awal yang akan dilakukan adalah penentuan terhadap topik yang akan diangkat dalam penelitian. Tahap ini merupakan tahapan yang sangat penting karena akan menentukan arah dari keseluruhan proses penelitian dan berisi topik permasalahan yang akan diangkat untuk selanjutnya diteliti. Pemilihan topik juga didasarkan pada beberapa faktor seperti ketersediaan data, minat peneliti serta kontribusi hasil terhadap masalah yang ada pada dunia nyata.

Selanjutnya akan dilakukan pengumpulan, peninjauan, serta analisa terhadap berbagai sumber ilmiah yang relevan dengan topik penelitian melalui studi literatur. Pada tahap ini peneliti mencari sumber seperti jurnal, buku, dan artikel untuk memahami konsep serta teori untuk dijadikan sebagai bahan kutipan sesuai topik penelitian. Hasil studi menunjukkan bahwa *Long – Short term memory* (LSTM) memiliki performa yang lebih baik dibandingkan beberapa algoritma lain seperti XGboost, GRU, maupun ARIMA dan memiliki kemampuan yang baik dalam menangani data deret waktu, namun masih memiliki keterbatasan dalam pemilihan parameter optimal. Oleh karena itu, penelitian ini mengusulkan integrasi antara algoritma *Long-Short-Term Memory* (LSTM) dengan algoritma NADAM untuk mengoptimalkan performa model.

Penggunaan algoritma *Artificial Intelligence* (AI) memang semakin banyak digunakan untuk menghadapi berbagai permasalahan yang ada di dunia nyata akibat perkembangannya yang pesat, salah satu contoh penerapannya adalah untuk memprediksi kekuatan magnitudo gempa bumi. Dalam penelitian ini, Algoritma NADAM akan digunakan untuk membantu mengoptimalkan algoritma *Long-Short-Term Memory* (LSTM) sehingga dapat menghasilkan model prediksi yang optimal.

Berdasarkan studi literatur yang dilakukan, penelitian ini merumuskan permasalahan utama terkait dengan keterbatasan dalam pemilihan parameter optimal pada model yang akan dilatih untuk memprediksi kekuatan magnitudo gempa bumi. Tujuan dari penelitian ini adalah untuk meningkatkan performa algoritma LSTM dengan menggunakan algoritma NADAM Sehingga dapat menghasilkan model prediksi kekuatan magnitudo gempa yang lebih akurat dan stabil.

2.3 Optimasi LSTM dengan NADAM untuk Meningkatkan Performa

Penelitian ini menggunakan data terkait kejadian gempa bumi yang diambil dari website <https://repogempa.bmkg.go.id/> Badan Meteorologi Klimatologi dan Geofisika (BMKG) yang berjumlah 13420 kasus data gempa bumi di pulau sumatera sejak november 2008 – april 2023. Kolom data awal yang diperoleh melalui laman BMKG pada saat penarikan data dapat dilihat pada Tabel 1 berikut.

Tabel 1. Kolom Pada *Dataset*

Nama Kolom	Keterangan
No	Nomor kejadian gempa
Event ID	Id kejadian gempa
Date time	Tanggal, bulan, tahun, dan waktu kejadian gempa
Latitude	Titik Garis lintang kejadian gempa
Longitude	Titik garis bujur kejadian gempa
Magnitude	Nilai kekuatan magnitudo gempa yang terjadi
Mag Type	Skala pengukuran kekuatan gempa
Depth (km)	Kedalaman pusat gempa dari permukaan bumi
Phase Count	Jumlah gelombang seismik yang terdeteksi
Azimuth Gap	Celah sudut terbesar antara stasiun-stasiun seismik yang menangkap gempa
Location	Lokasi kejadian gempa
Agency	Lembaga yang melakukan pencatatan kejadian gempa

Dari dataset yang diperoleh seperti pada Tabel 1 diatas, variabel yang nantinya akan digunakan pada penelitian ini adalah *datetime*, *latitude*, *longitude*, *magnitude*, dan *depth* (km).

Sebelum data pada *dataset* digunakan untuk melatih model, terlebih dahulu akan dilakukan *preprocessing* pada data yang dimiliki. Tahap *preprocessing* berisi kegiatan mengintegrasikan data, pembersihan data dari nilai yang kurang relevan dan

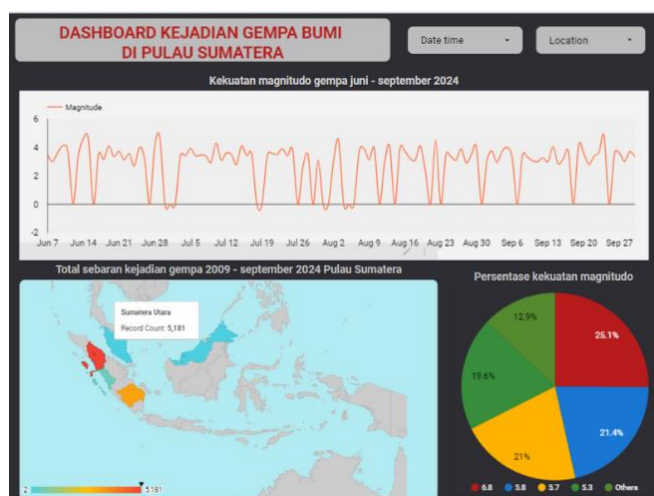
transformasi data yang dilakukan pada tahap awal penelitian untuk meningkatkan tingkat akurasi hasil penelitian [4]. Selanjutnya akan dilakukan *splitting* pada *dataset* yang dimiliki. Pada tahap ini *dataset* akan dilakukan pembagian dengan tujuan untuk membantu model yang akan dibuat dalam menemukan hubungan antar fitur sehingga dapat memprediksi kemungkinan yang terjadi dimasa mendatang. Selain itu, data akan dinormalisasikan melalui proses *Feature engineering* sehingga menciptakan representasi data yang sesuai untuk membuat model *machine learning* [5]. Beberapa teknik *feature engineering* diantaranya adalah *encoding variabel*, normalisasi data menjadi skala yang konsisten serta pembuatan fitur baru dari data yang telah ada.

Pada tahapan ini, metode yang dipilih untuk melakukan prediksi ialah menggunakan algoritma *Long Short Term Memory* (LSTM). Adapun langkah penggunaan algoritma LSTM dengan NADAM adalah sebagai berikut :

- a. Import *library* keras untuk membuat model LSTM, kemudian inisialisasi parameter dengan menentukan banyak layer pada LSTM, *learning rate*, jumlah *epoch* dan *time step* yang akan digunakan pada pembuatan model LSTM kemudian jenis optimasi yang akan digunakan untuk mempercepat konvergensi dari model, dalam hal ini NADAM akan digunakan sebagai optimasi karena dianggap lebih cepat dalam beberapa kasus dibandingkan dengan algoritma optimasi lainnya. algoritma ini akan bekerja dengan memperbarui bobot untuk menghaluskan gradien yang berdampak untuk mempercepat proses pelatihan dan peningkatan akurasi model.
- b. Kemudian, untuk melatih model akan digunakan dataset yang telah dibagi menjadi data *training* dan *testing*.
- c. Selanjutnya algoritma LSTM yang dioptimasi dengan NADAM akan dilatih dengan data yang telah diatur dalam *time step* dengan setiap *time step* dari urutan input akan dilatih untuk memprediksi nilai berikutnya. Jika *time step* yang ditentukan misalnya adalah 5, maka setiap *time step* akan terdiri dari lima input berurutan dengan target prediksi adalah nilai urutan selanjutnya. Pada proses ini akan melibatkan semua gerbang LSTM seperti *forget gate*, *input gate*, *cell state*, dan *output gate* yang masing – masing berperan untuk mengatur berapa banyak informasi dari langkah sebelumnya yang dilupakan, memperbarui informasi, menyimpan informasi baru, dan memberikan keluaran dari data yang diperbarui.
- d. Setelah model berhasil dibangun dan dilatih, masukkan data uji ke model yang telah dirancang dengan juga menentukan *time step* seperti pelatihan.
- e. Setelah hasil prediksi didapatkan, kemudian dapat melakukan denormalisasi data yang sebelumnya berada di rentang 0 – 1 menjadi bentuk nilai aslinya Sehingga data dapat diinterpretasikan dalam bentuk nyata.

Setelah model berhasil dilatih dan digunakan untuk memprediksi, tahap selanjutnya akan dilakukan uji kualitas prediksi yang telah dihasilkan oleh model dengan menggunakan beberapa algoritma untuk mengevaluasi model prediksi yang dibuat seperti *Mean Absolut Error* (MAE), *Mean Absolut Percentage Error* (MAPE), *Mean Square Error* (MSE) dan *Root Mean Square Error* (RMSE). Hal ini dilakukan untuk mengetahui seberapa akurat kinerja model dalam memprediksi pada saat training maupun testing data, sehingga kemudian model dapat digunakan untuk melakukan prediksi kejadian dimasa depan.

Setelah dilakukan evaluasi terhadap hasil prediksi model, selanjutnya hasil prediksi tersebut akan ditransformasikan kedalam bentuk sebuah visualisasi. Visualisasi merupakan suatu cara untuk menggambarkan suatu informasi secara grafis sehingga membantu mempermudah dalam pemahaman dan analisa data [6]. Pada tahapan ini, akan dilakukan penggambaran dari data hasil prediksi menggunakan bantuan tools *Google Looker Studio* sehingga diharapkan dapat membantu dalam mempermudah memahami hasil prediksi yang telah dilakukan. Contoh hasil visualisasi hasil prediksi yang diharapkan dapat dilihat pada Gambar 2 berikut ini.



Gambar 2 . Contoh Visualisasi Hasil Prediksi

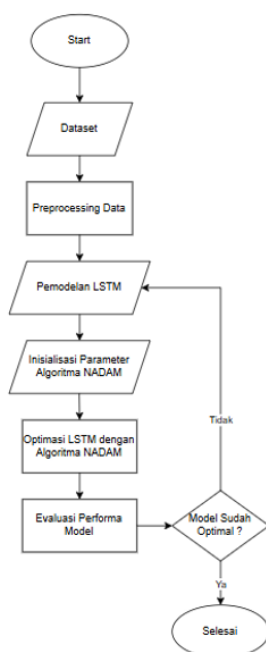
Pada *dashboard* contoh hasil visualisasi prediksi gempa seperti yang ditampilkan pada Gambar 2, terdapat beberapa informasi penting yang dapat dimanfaatkan oleh masyarakat maupun pihak terkait kebencanaan di wilayah sumatera. Informasi yang dimuat dalam *dashboard* tersebut mencakup prediksi kekuatan magnitudo gempa yang diperkirakan terjadi selama 1 bulan mendatang. Visualisasi ini akan menampilkan hasil prediksi kekuatan gempa setiap harinya di masing-masing wilayah di pulau sumatera yang memberikan gambaran mengenai potensi tingkat kekuatan bencana gempa di masa mendatang.

Selain itu, *dashboard* ini juga memuat visualisasi jumlah kejadian gempa yang terjadi di masing-masing wilayah di pulau sumatera dalam periode tertentu beserta grafik persentase prediksi kekuatan magnitudo gempa yang terjadi. Dengan visualisasi yang terintegrasi ini, diharapkan dapat membantu masyarakat dan lembaga kebencanaan daerah setempat dalam meningkatkan kewaspadaan terhadap kemungkinan kejadian gempa yang akan datang sehingga lebih siap dalam menghadapi ancaman gempa yang mungkin terjadi.

3. HASIL DAN PEMBAHASAN

Pada bagian ini akan disajikan hasil dari penelitian mengenai optimasi LSTM dengan menggunakan NADAM untuk melakukan prediksi terhadap kekuatan magnitudo gempa dipulau sumatera. Selain itu, pembahasan juga akan mencakup evaluasi dari kinerja model menggunakan metrik evaluasi seperti MAE, MAPE, MSE, RMSE serta visualisasi hasil prediksi untuk memberikan pemahaman yang lebih baik mengenai efektivitas dari model yang telah dikembangkan.

3.1 Alur Optimasi LSTM dengan NADAM Untuk Meningkatkan Performa



Gambar 3. Alur Penelitian LSTM dengan Optimasi NADAM

Gambar 3 diatas menjelaskan alur kerja penelitian yang memadukan antara algoritma *Long Short Term Memory* (LSTM) dengan Algoritma *Nesterov-accelerated Adaptive Moment Estimation* (NADAM) untuk optimasi model. Proses diawali dengan tahapan preprocessing pada data yang meliputi beberapa tahapan diantaranya adalah memfilter secara manual wilayah yang hanya termasuk ke dalam pulau Sumatra saja dengan menggunakan bantuan dari Microsoft Excel sebelum dilakukan tahap *preprocessing* selanjutnya. Sehingga total data yang diperoleh berjumlah 13419 baris data yang sebelumnya berjumlah 100239 pada *dataset* awal yang diperoleh. Selain itu, akan dilakukan pula beberapa tahapan *preprocessing* lainnya yang meliputi pengecekan duplikat pada data, penyesuaian tipe data pada masing – masing kolom yang ada pada *dataset*, serta memilih kolom apa saja yang ada pada *dataset* yang akan digunakan untuk melakukan peramalan.

a. Data Duplikat

Duplikat data adalah nilai yang sama yang muncul lebih dari 1 kali dalam suatu dataset. Pada *dataset* kejadian gempa yang digunakan, tidak ditemukan adanya data duplikat setelah dilakukan pengecekan. Sehingga dapat dipastikan bahwa data yang digunakan bebas dari duplikasi dan siap untuk dilakukan proses *cleaning* lebih lanjut.

b. Penyesuaian Tipe Data

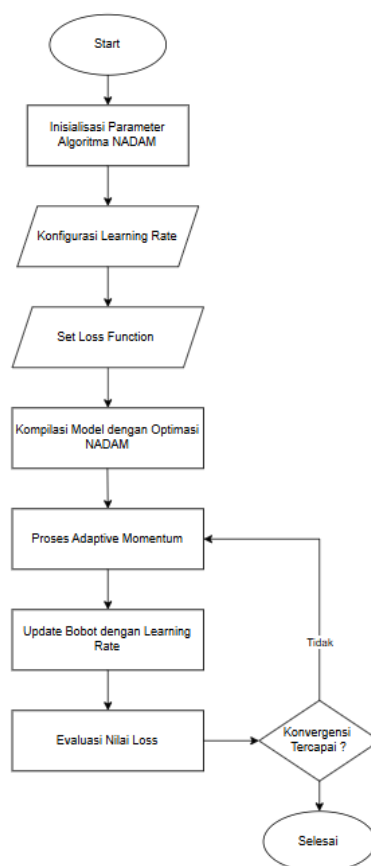
Pada penelitian, Tipe data pada kolom Date time akan dirubah menjadi tipe data *datetime* sesuai dengan data yang dimuat oleh kolom tersebut. Proses ini akan menggunakan fungsi `pd.to_datetime()` yang terdapat pada *library* pandas pada python.

c. Pemilihan Kolom Data

Pada penelitian ini, attribut yang akan digunakan meliputi kolom Date time yang di ubah menjadi tahun, bulan, hari, jam, menit, detik untuk membantu model yang akan dilatih dalam memahami pola kejadian gempa yang terjadi.

Setelah tahapan *preprocessing* selesai, selanjutnya data akan dibagi menjadi data latih (*data train*) dan data uji (*data test*) untuk digunakan dalam proses pembelajaran dan pengujian. Pada tahap optimasi, algoritma NADAM digunakan untuk membantu mempercepat konvergensi model pada tahap pelatihan dengan menyesuaikan bobot secara adaptif, sehingga model dapat belajar lebih stabil dan efisien. Hasil keluaran dari model kemudian akan dilakukan evaluasi menggunakan beberapa metrik evaluasi yang telah disebutkan sebelumnya dengan tujuan untuk mengukur sejauh mana kemampuan model dalam belajar dan memprediksi.

3.1.1 Melakukan Optimasi LSTM dengan NADAM



Gambar 4. Alur Kerja Optimasi NADAM

Berdasarkan Gambar 4 diatas, model yang akan dilatih nantinya akan digunakan untuk memprediksi kekuatan magnitudo gempa. Sehingga untuk mengoptimalkan performa model, dilakukan penyesuaian *hyperparameter* menggunakan algoritma optimasi NADAM. Proses ini mencakup konfigurasi *learning rate*, penetapan *loss function*, serta pembaruan bobot dengan *adaptive momentum*. Penyesuaian dilakukan secara iteratif hingga model mencapai konvergensi sehingga diharapkan dapat memberikan akurasi yang lebih optimal. Algoritma NADAM yang akan digunakan untuk melakukan optimasi pada model yang akan dilatih dapat dilihat pada Tabel 2 berikut ini.

Tabel 2. Pseudocode Algoritma NADAM

```

Initialize parameters:
    m = 0 # First moment vector
    v = 0 # Second moment vector
     $\mu = 0$  # Nesterov momentum
    t = 0 # Time step
     $\beta_1 = 0.9$  # Decay rate for first moment
     $\beta_2 = 0.999$  # Decay rate for second moment
     $\epsilon = 1e-7$  # Small constant for numerical stability
     $\eta$  = learning rate # Initial learning rate

For each iteration:
    t = t + 1
    Compute gradient g of the loss function with respect to parameters  $\theta$ 

    # Update first moment estimate
    m =  $\beta_1 * m + (1 - \beta_1) * g$ 

    # Update second moment estimate
    v =  $\beta_2 * v + (1 - \beta_2) * g^2$ 

    # Compute bias-corrected estimates
    m_hat = m / (1 -  $\beta_1^t$ )
    v_hat = v / (1 -  $\beta_2^t$ )

    # Compute Nesterov momentum term
     $\mu = \beta_1 * m\_hat + (1 - \beta_1) * g / (1 - \beta_1^t)$ 

    # Update parameters
     $\theta = \theta - \eta * \mu / (\text{sqrt}(v\_hat) + \epsilon)$ 

Return updated parameters  $\theta$ 

```

Setelah dilakukan eksperimen untuk mencari parameter terbaik dari model LSTM dengan optimasi NADAM, arsitektur dari model *final* LSTM – NADAM yang diperoleh yang nantinya akan digunakan untuk melakukan prediksi dapat dilihat pada Tabel 3 dan Tabel 4 berikut ini.

Tabel 3. Pseudocode LSTM - NADAM

Inisialisasi model sebagai model Sequential.

Tambahkan layer pertama:

- a. Layer LSTM dengan 128 unit.
- b. Gunakan aktivasi 'tanh'.
- c. Set return_sequences=True agar mengembalikan seluruh urutan output.
- d. Tetapkan input_shape dengan dimensi (jumlah waktu langkah, fitur per langkah).

Tambahkan layer Dropout dengan rasio 0.3 untuk mencegah overfitting.

Tambahkan layer kedua:

- a. Layer LSTM dengan 64 unit.
- b. Gunakan aktivasi 'tanh'.
- c. Set return_sequences=False untuk hanya mengambil output terakhir.

Tambahkan layer Dropout dengan rasio 0.2.

Tambahkan layer Dense dengan:

- a. 16 unit neuron.
- b. Aktivasi 'relu'.
- c. Regularisasi L2 dengan faktor 0.00004.

Tambahkan layer Dense dengan:

- a. 1 unit output.
- b. Aktivasi 'linear' untuk prediksi nilai kontinu.

Kompilasi model dengan:

- a. Optimizer Nadam dengan learning rate 0.00003.
- b. Loss function Mean Squared Error (MSE).
- c. Metrics 'mae' (Mean Absolute Error).

Tampilkan ringkasan arsitektur model.

Tabel 4. Model Final Arsitektur LSTM – NADAM

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 1000, 128)	70,656
dropout (Dropout)	(None, 1000, 128)	0

lstm_1 (LSTM)	(None, 64)	49,408
dropout_1 (Dropout)	(None, 64)	0
dense (Dense)	(None, 16)	1,040
dense_1 (Dense)	(None, 1)	17
Total params		121,121
Trainable params		121,121
Non-trainable params		0

Tabel 3 dan Tabel 4 diatas menunjukkan arsitektur model yang didapat dengan NADAM terdiri dari 2 layer LSTM dengan *layer* pertama memiliki 128 unit yang kemudian menghasilkan *output shape* (Non, 1000, 128) serta memiliki 70,656 parameter lalu diikuti dengan layer dropout untuk mencegah terjadinya *overfitting*. Selain itu terdapat pula *layer* kedua LSTM dengan 64 unit yang menghasilkan *output shape* (None,64) serta memiliki 49,408 parameter lalu diikuti dengan *layer dropout*. Model ini akan dilakukan pelatihan dengan menggunakan data latih lalu kemudian model akan melakukan prediksi kekuatan gempa dengan menggunakan data pengujian. Hasil prediksi akan di normalisasikan kembali ke skala yang asli lalu kemudian akan dilakukan evaluasi menggunakan MAE, MAPE, MSE, dan RMSE.

3.1.2 Evaluasi Model Menggunakan MAE, MAPE, MSE, dan RMSE

Evaluasi dilakukan untuk menilai sejauh mana model yang telah dilatih mampu melakukan peramalan dengan baik. Tahap evaluasi diterapkan pada data pengujian dengan pembagian dataset yang berbeda yaitu 90 :10, 80 : 20 , dan 70 : 30 dan menerapkan pemberian *time step* yang bervariasi meliputi timestep 7, 60, 120, 250, 500, dan 1000. Hasil evaluasi pengujian dan pemberian *timestep* pada masing – masing data ditunjukkan pada Tabel 5 dibawah ini.

Tabel 5. Hasil Pengujian Pada Pembagian Dataset dan Timestep

Pembagian Dataset	Time Step	MAE	MAPE	MSE	RMSE
90 : 10	7	0.5353	15.74%	0.4794	0.6924
90 : 10	60	0.535	15.75%	0.478	0.6914
90 : 10	120	0.5372	15.79%	0.4792	0.6891
90 : 10	250	0.561	16.30%	0.5256	0.725
90 : 10	500	0.555	15.88%	0.4599	0.6707
90 : 10	1000	0.5751	15.93%	0.4615	0.6794
80 : 20	7	0.565	18.50%	0.478	0.69
80 : 20	60	0.578	16.62%	0.4925	0.7017
80 : 20	120	0.6211	20.99%	0.551	0.7423
80 : 20	250	0.5367	16.28%	0.4698	0.6854
80 : 20	500	0.5296	16.33%	0.4703	0.6858

80 : 20	1000	0.5269	15.81%	0.4586	0.6772
70 : 30	7	0.54	16.20%	0.47	0.685
70 : 30	60	0.6621	23.37%	0.6452	0.8033
70 : 30	120	0.73	25.50%	0.835	0.912
70 : 30	250	0.776	28.38%	0.921	0.959
70 : 30	500	0.6854	23.48%	0.7464	0.8639
70 : 30	1000	0.6577	22.70%	0.6648	0.8154

Dari hasil Tabel 5 diatas, dapat diketahui bahwa dari beberapa kombinasi yang dilakukan, kombinasi yang baik diperoleh dengan menggunakan 80% data latih dan 10% data uji serta 1000 *timestep* dengan hasil MAE, MAPE, MSE, dan RMSE yang lebih rendah diantara kombinasi yang lainnya.

3.2 Analisa dan Validasi

3.2.1 Analisa Terhadap Hasil Prediksi Yang Dihasilkan Oleh model

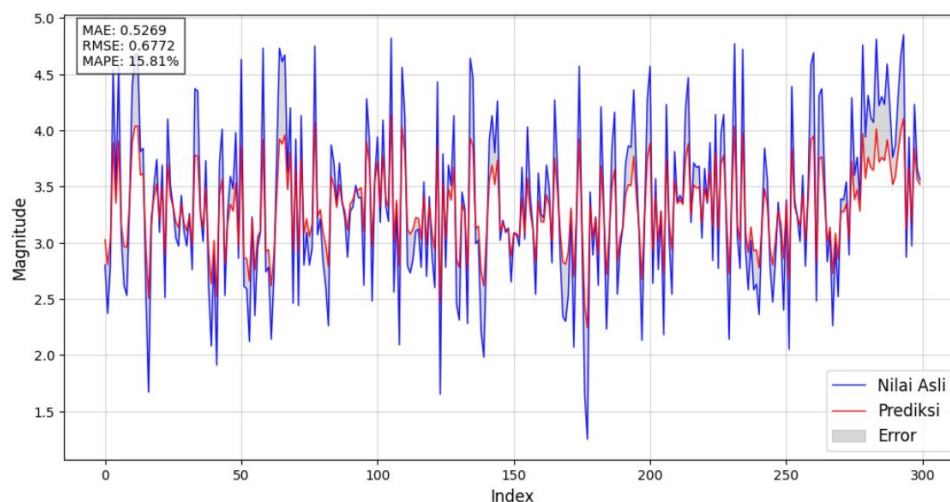
Pada tahapan selanjutnya akan disajikan hasil dari jumlah *epoch* serta *batch size* yang diterapkan pada model. *Epoch* yang akan diuji akan berada di rentang 50 – 200 serta *batch size* 32, 50, 60, dan 64 untuk mencari hasil yang paling optimal.

Tabel 6. Hasil Uji Model Pada Timestep Terbaik

Algoritma	Batch Size	Epoch	MAE	MAPE	MSE	RMSE
LSTM - NADAM	32	50	0.5280	16.50%	0.4819	0.6942
LSTM			0.5915	18.50%	0.5406	0.7352
LSTM - NADAM	32	60	0.5269	15.81%	0.4586	0.6772
LSTM			0.6023	18.85%	0.5549	0.7449
LSTM - NADAM	32	100	0.5697	16.79%	0.5332	0.7302
LSTM			0.6375	19.88%	0.6282	0.7926
LSTM - NADAM	32	150	0.5336	16.08%	0.4676	0.6838
LSTM			0.6707	21.09%	0.6855	0.8279
LSTM – NADAM	32	200	0.5406	16.63%	0.4646	0.6816
LSTM			0.6433	20.42%	0.6326	0.7954
LSTM – NADAM	50	50	0.5390	16.92%	0.4520	0.6723
LSTM			0.5841	18.47%	0.5213	0.7220
LSTM – NADAM	50	60	0.5345	16.28%	0.4597	0.6780
LSTM			0.6015	19.28%	0.5505	0.7420
LSTM - NADAM	50	100	0.5325	15.98%	0.4677	0.6839
LSTM			0.6167	18.34%	0.5859	0.7655
LSTM – NADAM	50	150	0.5432	16.48%	0.4731	0.6878
LSTM			0.6557	20.42%	0.6648	0.8153
LSTM – NADAM	50	200	0.5346	16.65%	0.5321	0.7294
LSTM			0.6635	20.92%	0.6875	0.8292
LSTM – NADAM	60	50	0.5351	16.86%	0.4632	0.6806
LSTM			0.5795	18.58%	0.5128	0.7161
LSTM – NADAM	60	60	0.5366	15.91%	0.4803	0.6931
LSTM			0.5900	18.30%	0.5437	0.7374
LSTM – NADAM			0.5304	15.88%	0.4615	0.6794

LSTM	60	100	0.6171	19.37%	0.5951	0.7714
LSTM - NADAM	60	150	0.5373	16.24%	0.4703	0.6858
LSTM			0.6495	20.25%	0.6605	0.8127
LSTM – NADAM	60	200	0.549	16.54%	0.4864	0.6974
LSTM			0.6420	20.05%	0.6371	0.7982
LSTM – NADAM	64	50	0.5289	16.51%	0.5195	0.7208
LSTM			0.5898	18.79%	0.5340	0.7307
LSTM – NADAM	64	60	0.5237	16.21%	0.4712	0.6864
LSTM			0.5870	19.00%	0.5259	0.7252
LSTM – NADAM	64	100	0.5556	16.85%	0.5068	0.7119
LSTM			0.6215	19.82%	0.5973	0.7728
LSTM - NADAM	64	150	0.5292	16.26%	0.4615	0.6858
LSTM			0.6518	20.75%	0.6570	0.8106
LSTM – NADAM	64	200	0.5578	17.10%	0.4904	0.7003
LSTM			0.6373	20.26%	0.6299	0.7937

Dari Tabel 6 diatas menunjukkan bahwa LSTM yang dioptimasi dengan NADAM memperoleh performa terbaiknya pada epoch 60 dan *batch size* 32. Nilai MAE, MAPE, MSE, dan RMSE rendah yang diperoleh menunjukkan bahwa prediksi yang dihasilkan lebih akurat dibandingkan dengan kombinasi lainnya dalam memprediksi kekuatan magnitudo gempa yang terjadi di pulau sumatera. Grafik yang menunjukkan perbandingan antara data prediksi dan data aktual magnitudo gempa pada *epoch* dan *batch size* terbaik yang diperoleh ditunjukkan pada Gambar 5 berikut ini.

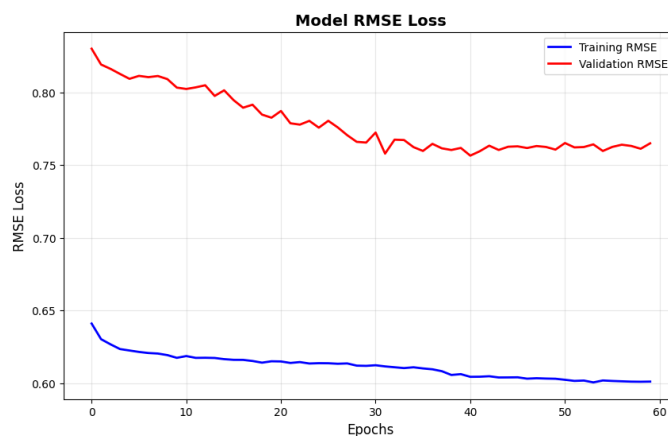


Gambar 5. Grafik Perbandingan Data Prediksi dan Data Aktual

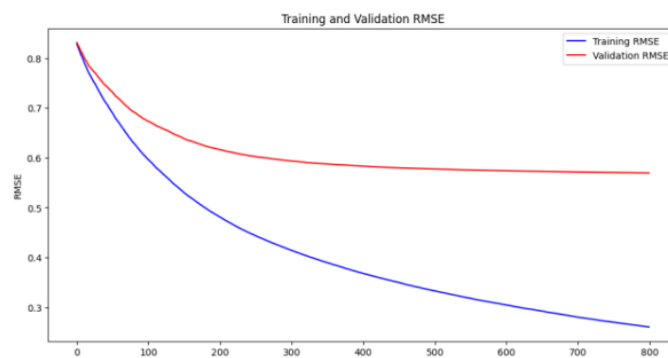
Pada Gambar 5 diatas, terdapat garis biru yang menggambarkan data aktual dari kekuatan magnitudo gempa berdasarkan dataset yang digunakan dalam penelitian ini. Sementara itu, garis berwarna merah menunjukkan nilai prediksi yang dihasilkan oleh model yang telah melalui proses pelatihan. Area berwarna abu-abu, yang terletak di antara garis biru dan garis merah, merepresentasikan gap atau selisih antara nilai prediksi dan nilai aktual dari kekuatan magnitudo gempa.

Semakin kecil gap yang terdapat di antara kedua garis tersebut, maka semakin baik pula kinerja model dalam melakukan prediksi, karena menunjukkan bahwa hasil prediksi memiliki tingkat akurasi yang tinggi terhadap data yang sebenarnya. Sebaliknya, jika gap yang terbentuk cukup besar, maka dapat disimpulkan bahwa model masih memiliki tingkat kesalahan yang cukup tinggi dalam melakukan estimasi terhadap magnitudo gempa yang terjadi.

Dari hasil evaluasi yang dilakukan pada model, dapat diketahui bahwa performa model LSTM – NADAM berada di bawah penelitian sebelumnya yang menggunakan algoritma XGboost untuk memprediksi kekuatan magnitudo gempa. Hasil penelitian sebelumnya menunjukkan bahwa algoritma XGboost memperoleh nilai MSE sebesar 0.1 pada data pengujian [7]. Sedangkan pada algoritma LSTM – NADAM memperoleh nilai MSE sebesar 0.4 pada saat dilakukan pengujian model. Berikut adalah contoh grafik perbandingan nilai loss pada penelitian dengan menggunakan LSTM – NADAM dengan penelitian sebelumnya menggunakan algoritma XGboost yang dapat dilihat pada Gambar 6 dan Gambar 7 berikut ini.



Gambar 6. Nilai Loss LSTM – NADAM



Gambar 7. Nilai Loss XGboost Pada Penelitian Sebelumnya [7]

Berdasarkan kedua grafik pada Gambar 6 dan Gambar 7 diatas, dapat diketahui bahwa model LSTM dengan optimasi NADAM menunjukkan tren penurunan RMSE baik pada data training maupun validation yang menandakan bahwa model memiliki kemampuan yang semakin baik dalam memahami pola pada data.

Namun, terlihat pula bahwa nilai validation RMSE tetap lebih tinggi dibandingkan training RMSE dan cenderung stagnan setelah beberapa epoch terakhir. Hal ini menunjukkan bahwa model LSTM kemungkinan terlalu menyesuaikan dengan data training sehingga sedikit kurang mampu dalam melakukan generalisasi dengan baik pada data baru.

Sementara itu, pada grafik model XGBoost menunjukkan performa yang lebih stabil dengan tren penurunan RMSE yang konsisten pada training dan validation. Perbedaan antara training RMSE dan validation RMSE yang tidak terlalu besar menandakan bahwa model ini memiliki kemampuan generalisasi yang lebih baik dibandingkan LSTM dalam konfigurasi yang digunakan.

3.2.2 Validasi Keakuratan Model Prediksi Kekuatan Magnitudo Gempa

Selanjutnya hasil performa model dalam melakukan prediksi dapat dianalisis lebih lanjut dengan membandingkan nilai-nilai prediksi dengan nilai aktual yang tersaji dalam Tabel 7. Melalui tabel tersebut, dapat dilihat secara lebih rinci sejauh mana perbedaan antara hasil prediksi yang dihasilkan oleh model dengan nilai sebenarnya yang ada pada dataset. Hal ini penting untuk mengevaluasi tingkat keakuratan dan keandalan model dalam memprediksi kejadian gempa berdasarkan pola data yang telah dipelajari selama proses pelatihan.

Tabel 7. Nilai Prediksi dan Aktual Kekuatan Magnitudo Gempa

Aktual	Prediksi
3.4	3.2
3.04	3.2
3.3	3.4
3.5	3.3
3.1	3.2
2.8	3.07
3.6	3.3

Tabel 7 diatas menunjukkan perbandingan antara nilai aktual dan hasil prediksi kekuatan magnitudo gempa. Dengan menggunakan model LSTM yang dioptimasi menggunakan algoritma NADAM, selisih atau perbedaan antara nilai aktual dan prediksi yang didapat menunjukkan hasil yang cukup akurat, dengan tingkat perbedaan yang relatif kecil. Hal ini mengindikasikan bahwa model LSTM dengan optimasi NADAM memiliki kemampuan yang baik dalam memprediksi kekuatan magnitudo gempa.

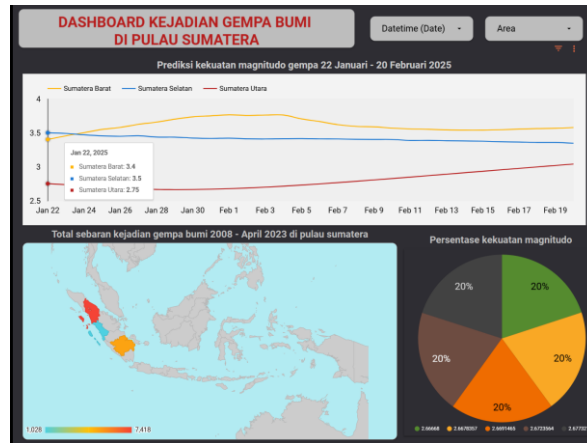
Setelah model berhasil untuk dilatih dan diuji, model kemudian akan digunakan untuk memprediksi terhadap kemungkinan kekuatan magnitude gempa yang terjadi selama 1 bulan mendatang di pulau sumatera. Prediksi akan dilakukan pada tiap bagian wilayah yang ada di pulau sumatera meliputi sumatera bagian utara, sumatera bagian barat, dan sumatera bagian selatan. Contoh hasil prediksi model yang telah dilatih untuk memprediksi kekuatan magnitudo gempa dapat dilihat pada Tabel 8 berikut.

Tabel 8. Hasil Prediksi Model LSTM - NADAM

Day	Predicted Magnitude	Area	Datetime
1	2.754127	Sumatera Utara	1/22/2025
2	2.7397594	Sumatera Utara	1/23/2025
3	2.72216	Sumatera Utara	1/24/2025
4	2.7046945	Sumatera Utara	1/25/2025
5	2.6893754	Sumatera Utara	1/26/2025
6	2.677321	Sumatera Utara	1/27/2025
7	2.6691465	Sumatera Utara	1/28/2025
8	2.66668	Sumatera Utara	1/29/2025
9	2.6678357	Sumatera Utara	1/30/2025
10	2.6723564	Sumatera Utara	1/31/2025
1	3.5024002	Sumatera Selatan	1/22/2025
2	3.4928267	Sumatera Selatan	1/23/2025
3	3.4726686	Sumatera Selatan	1/24/2025
4	3.4572494	Sumatera Selatan	1/25/2025
5	3.451121	Sumatera Selatan	1/26/2025
6	3.4589512	Sumatera Selatan	1/27/2025
7	3.4395554	Sumatera Selatan	1/28/2025
8	3.4378335	Sumatera Selatan	1/29/2025
9	3.4249275	Sumatera Selatan	1/30/2025
10	3.4194088	Sumatera Selatan	1/31/2025
1	3.4045222	Sumatera Barat	1/22/2025
2	3.4572284	Sumatera Barat	1/23/2025
3	3.5033627	Sumatera Barat	1/24/2025
4	3.5520513	Sumatera Barat	1/25/2025
5	3.5817838	Sumatera Barat	1/26/2025
6	3.6263008	Sumatera Barat	1/27/2025
7	3.6597369	Sumatera Barat	1/28/2025
8	3.7024279	Sumatera Barat	1/29/2025
9	3.7372503	Sumatera Barat	1/30/2025
10	3.7524805	Sumatera Barat	1/31/2025

Dari Tabel 8 diatas, dapat dilihat bahwa model yang telah dilatih dan diuji pada tahap sebelumnya mampu digunakan untuk melakukan prediksi terhadap kemungkinan kekuatan gempa yang terjadi dimasa mendatang. hasil model memberikan 30 frekuensi gempa yang terjadi di setiap wilayah bagian dipulau sumatera dengan kekuatan magnitudo gempa yang bervariasi

Setelah model berhasil digunakan melakukan prediksi, Hasil prediksi yang diperoleh kemudian akan divisualisasikan menggunakan *tools google looker studio* untuk memudahkan dalam memahami hasil prediksi yang dihasilkan. Hasil visualisasi prediksi menggunakan *tools looker studio* dapat dilihat pada Gambar 8 berikut.



Gambar 8. Visualisasi Hasil Prediksi

Dari hasil visualisasi prediksi kekuatan magnitudo gempa pada Gambar 8 diatas, dapat diketahui bahwa hasil prediksi model memberikan nilai yang bervariasi terhadap kekuatan magnitudo yang diperkirakan terjadi di masa mendatang di pulau sumatera. Visualisasi hasil prediksi menunjukkan bahwa kekuatan magnitudo gempa dengan kekuatan yang cukup tinggi terjadi di pulau sumatera bagian barat dengan rata – rata kekuatan magnitudo yang terjadi satu bulan kedepan sebesar 3.6 magnitudo sedangkan kekuatan magnitudo gempa yang paling rendah terjadi di pulau sumatera bagian utara dengan rata – rata kekuatan magnitudo gempa sebesar 2.79 magnitudo. namun meskipun demikian, jumlah kejadian gempa bumi yang terjadi di pulau sumatera bagian utara cenderung lebih banyak dibandingkan kejadian di pulau sumatera bagian lainnya dengan jumlah kasus dalam rentang waktu november 2008 – april 2023 sebesar 7418 kejadian diikuti dengan sumatera bagian selatan sebanyak 4973 kejadian dan sumatera bagian barat sebanyak 1028 kejadian.

Dengan memahami pola distribusi dan kekuatan gempa di pulau sumatera, diharapkan dapat membantu masyarakat maupun lembaga kebencanaan setempat dalam meningkatkan kesiapsiagaan dan strategi mitigasi guna meminimalkan dampak yang ditimbulkan akibat bencana gempa bumi yang mungkin terjadi di masa mendatang baik melalui edukasi masyarakat maupun penguatan infrastruktur pada wilayah yang rawan terkena dampak kejadian gempa bumi dipulau sumatera.

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa model LSTM yang dioptimasi dengan NADAM memiliki performa yang berada di bawah performa algoritma XGboost dalam memprediksi kekuatan magnitudo gempa. Namun meskipun demikian, LTSM – NADAM mampu menghasilkan performa yang dikategorikan baik dalam memprediksi kekuatan magnitudo gempa. Hasil evaluasi menunjukkan bahwa model LSTM – NADAM dengan *epoch* 60 dan *batch size* 32 mendapatkan hasil yang terbaik. Secara keseluruhan, pengoptimasian LSTM dengan NADAM menunjukkan bahwa proses tersebut dapat digunakan untuk mendapatkan hasil peramalan yang optimal.

Berdasarkan hasil prediksi kekuatan magnitudo gempa selama satu bulan ke depan yang telah dilakukan, didapatkan hasil bahwa sumatera barat memiliki rata-rata kekuatan magnitudo tertinggi sebesar 3,6 magnitudo sedangkan sumatera utara memiliki rata-rata kekuatan magnitudo terendah sebesar 2,79 magnitudo, dan sumatera selatan memiliki rata-rata kekuatan magnitudo sebesar 3,4 magnitudo. Meskipun demikian, frekuensi gempa di sumatera utara adalah yang tertinggi dengan 7.418 kejadian sejak november 2008 - april 2023, diikuti oleh sumatera selatan dengan 4.973 kejadian, dan sumatera barat dengan 1.028 kejadian. Dengan adanya informasi ini diharapkan dapat membantu meningkatkan kewaspadaan dan upaya mitigasi oleh masyarakat serta lembaga kebencanaan setempat untuk mengurangi dampak bencana yang ditimbulkan akibat gempa bumi.

DAFTAR PUSTAKA

- [1] H. Tantyoko, D. Kartika Sari, and A. R. Wijaya, "PREDIKSI POTENSIAL GEMPA BUMI INDONESIA MENGGUNAKAN METODE RANDOM FOREST DAN FEATURE SELECTION," 2023. [Online]. Available: <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/indexHenriTantyoko><http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index>
- [2] N. Faridah and B. Sugiantoro, "Analisis Optimasi Pada Algoritma Long ShortTerm Memory Untuk Memprediksi Harga Saham," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 1, p. 575, Feb. 2023, doi: 10.30865/mib.v7i1.5421.
- [3] A. A. Nurhalizah and Y. Cahyana, "Model Prediksi Kekuatan Gempa Dengan Menggunakan Algoritma Linear Regression Dan Support Vector Regression (Studi Kasus BMKG)," no. 2, p. 41, 2024, [Online]. Available: <https://www.kaggle.com/datasets/kekavigi/earthquakes-in-ndonesia>.
- [4] T. Gori, A. Sunyoto, and H. Al Fatta, "Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 1, pp. 215–224, Feb. 2024, doi: 10.25126/jtiik.20241118074.
- [5] C. Herdian, A. Kamila, and I. G. Agung Musa Budidarma, "Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi," *Technologia : Jurnal Ilmiah*, vol. 15, no. 1, p. 93, Jan. 2024, doi: 10.31602/tji.v15i1.13457.
- [6] T. Maharani Informatika, "VISUALISASI DATA UNTUK PEMODELAN PREDIKTIF: METODE DAN ALAT."
- [7] A. Wibowo, "MESA : Jurnal Teknik PREDIKSI KEKUATAN GEMPA MENGGUNAKAN MACHINE LEARNING DENGAN MODEL XGBOOST SEBAGAI LANGKAH STRATEGIS DALAM PERENCANAAN STRUKTUR BANGUNAN TAHAN GEMPA DI INDONESIA".