



Peramalan Kadar PM₁₀ Menggunakan Algoritma Long Short-Term Memory (LSTM) Sebagai Acuan Ketersediaan Ruang Terbuka Hijau Di Kota Jambi

Ulfa Khaira^{1*}, Mutia Fadhila Putri¹, Shally Yanova²

¹Program Studi Informatika, Universitas Jambi, Indonesia.

²Program Studi Teknik Lingkungan, Universitas Jambi, Indonesia.

Artikel Info

Kata Kunci:

Long Short-Term Memory;
Peramalan;
PM10;
Polutan.

Keywords:

Long Short-Term Memory;
Forecasting;
PM10;
Pollutant.

Riwayat Artikel:

Submitted: 24 Januari 2025

Accepted: 30 Maret 2025

Published: 31 Maret 2025

Abstrak: Pada tahun 2023 jumlah kendaraan bermotor di Kota Jambi hampir mencapai 960.814 unit. Peningkatan volume lalu lintas akan menyebabkan peningkatan emisi polusi udara, yang dapat mengurangi kualitas udara. Salah satu polutannya adalah PM₁₀, yang berkontribusi terhadap polusi udara. *Particulate matter 10* (PM₁₀) adalah partikel materi yang berukuran kurang dari 10 mikrometer. PM₁₀ dapat berdampak negatif pada sistem pernapasan, seperti serangan asma, penurunan fungsi paru-paru, dan bahkan kematian. Salah satu solusi untuk mengatasi masalah polusi udara di Jambi adalah melalui pengembangan prediksi temporal kualitas udara menggunakan data historis. Dengan membangun model prediksi berdasarkan indeks polutan, kita dapat memproyeksikan kualitas udara secara harian. Penelitian ini bertujuan untuk membuat model prediksi konsentrasi PM₁₀ di Kota Jambi dengan metode Long Short Term Memory. Data yang digunakan adalah data konsentrasi PM₁₀ pada bulan Januari sampai Juni 2024 menggunakan algoritma jaringan saraf tiruan Long Short-Term Memory (LSTM). Penelitian ini dilakukan melalui beberapa tahapan yaitu pengumpulan data, praproses data, pembagian data, pembuatan model prediksi konsentrasi PM₁₀, dan MAPE. Penelitian ini menghasilkan model prediksi konsentrasi PM₁₀ dengan nilai RMSE sebesar 0,021 dan MAPE 0,11%, yang berarti model peramalan memiliki peramalan yang sangat akurat. Dengan kemampuannya memprediksi kadar PM₁₀ di masa depan berdasarkan pola historis. Informasi ini krusial dalam menentukan prioritas pengembangan RTH.

Abstract: In 2023, the number of motorized vehicles in Jambi City nearly reached 960,814 units. The increase in traffic volume will lead to a rise in air pollution emissions, which can reduce air quality. One of the pollutants is PM₁₀, which contributes to air pollution. Particulate matter 10 (PM₁₀) consists of material particles with a size of less than 10 micrometers. PM₁₀ can have negative impacts on the respiratory system, such as asthma attacks, decreased lung function, and even death. One solution to address the air pollution problem in Jambi is through the development of temporal air quality prediction using historical data. By building a prediction model based on pollutant indices, we can project daily air quality. This research aims to create a prediction model for PM₁₀ concentration in Jambi City using the Long Short-Term Memory method. The data used is PM₁₀ concentration data from January to June 2024, utilizing the Long Short-Term Memory (LSTM) artificial neural network algorithm. This research was conducted through several stages: data collection, data preprocessing, data splitting, PM₁₀ concentration prediction model creation, and MAPE calculation. This research resulted in a PM₁₀ concentration prediction model with an RMSE value of 0.021 and a MAPE of 0.11%, indicating that the forecasting model has very high accuracy. With its ability to predict future PM₁₀ levels based

on historical patterns, this information is crucial in determining the priorities for green open space (RTH) development.

Corresponding Author:

Ulfa Khaira

Email: ulfakhaira@unja.ac.id

PENDAHULUAN

Pencemaran udara dapat dijelaskan sebagai keberadaan zat-zat dalam udara pada tingkat konsentrasi yang cukup tinggi sehingga dapat menyebabkan gangguan pada manusia, hewan, dan tumbuhan (Maduna & Tomašić, 2017). Di Indonesia saat ini menggunakan Indeks Standar Pencemar Udara (ISPU) untuk mengukur kualitas udara, sesuai dengan Keputusan Menteri Negara Lingkungan Hidup Nomor 14 tahun 2020. Dalam peraturan tersebut, terdapat 7 parameter yang digunakan untuk menghitung ISPU, yaitu PM10, PM2.5, NO2, SO2, CO, O3, dan HC (Aji Tritamtama et al., 2023). Dari semua parameter yang ada dalam ISPU, PM10 merupakan parameter yang memiliki dampak negatif yang cukup signifikan dalam rentang yang relatif kecil. Particulate matter 10 (PM10) merupakan partikulat yang berukuran kurang dari 10 mikrometer (Gupta et al., 2020). PM10 berasal dari berbagai proses mekanis, seperti debu yang terangkat dari permukaan jalan, aktivitas industri, pertanian, serta bioaerosol (Vlasov et al., 2022). Meskipun tergolong partikel yang relatif besar, PM10 masih berpotensi terhirup hingga mencapai saluran pernapasan bagian atas, termasuk trachea, bronkus, dan alveolus di paru-paru, yang dapat memicu terjadinya infeksi saluran pernapasan atas (ISPA) (Priyankara et al., 2021).

Berdasarkan data Provinsi Jambi Dalam Angka 2024, jumlah kendaraan bermotor di Kota Jambi hampir mencapai angka satu juta unit. Angka tersebut mengalami lonjakan signifikan dari tahun sebelumnya, di mana pada tahun 2022 terdapat sekitar 740.623 unit kendaraan, sedangkan pada tahun 2023 mencapai 960.814 unit. Data ini memperlihatkan pertumbuhan yang pesat dalam jumlah kendaraan yang beredar di Kota Jambi dalam kurun waktu yang relatif singkat. Beberapa penelitian menunjukkan bahwa pertambahan volume lalu lintas juga akan mengakibatkan bertambahnya emisi polusi udara sehingga dapat menurunkan kualitas udara (Mangones et al., 2020; Wang et al., 2020). Salah satunya polutan gas CO dan PM10 yang menjadi kontributor polusi udara. Kondisi PM10 di beberapa titik di kota Jambi baik pada saat hari libur maupun hari kerja diketahui bahwa kawasan pinggir jalan raya masih merupakan yang tertinggi konsentrasi PM10-nya ($83,2 - 112,8 \mu\text{g}/\text{m}^3$), dibandingkan dengan kawasan perkantoran dan perumahan (Saputra et al., 2019).

Pemerintah Kota Jambi berusaha menjaga kualitas udara tetap baik dengan melakukan pemantauan secara teratur. Salah satu indikator yang dipantau adalah PM₁₀. Salah satu cara untuk mengatasi PM₁₀ yang terdapat pada udara ambien yaitu dengan memanfaatkan Ruang Terbuka Hijau (RTH) (Meo et al., 2021). Dalam upaya pemecahan masalah tingginya konsentrasi PM₁₀ udara ambien di area perkotaan, maka tanaman di area perkotaan lebih difungsikan sebagai penyerap polutan (Joshi et al., 2020). Tanaman di perkotaan identik dengan ruang terbuka hijau (RTH). Reduksi PM₁₀ bisa dilihat dari massa PM₁₀ yang menempel pada bagian tanaman. Semakin banyak tanaman, maka akan semakin banyak pula PM₁₀ yang terserap pada tanaman (Chiam et al., 2019).

Salah satu solusi untuk menangani isu pencemaran udara di Jambi adalah melalui pengembangan prediksi temporal terhadap kualitas udara dengan memanfaatkan data historis. Dengan membangun model prediksi berdasarkan indeks polutan, kita dapat memproyeksikan kualitas udara secara harian. Pada penelitian ini akan memprediksi kadar PM₁₀ di Kota Jambi menggunakan algoritma jaringan syaraf tiruan Long Short-Term Memory (LSTM). LSTM adalah model neural network yang efektif digunakan untuk melakukan prediksi time-series. Beberapa penelitian berhasil menunjukkan performa LSTM yang lebih baik jika dibandingkan dengan metode lain dalam melakukan prediksi time series (Huang et al., 2022; Song et al., 2020), performa LSTM berhasil mengungguli model

ARIMA (Latif et al., 2023; Siami-Namini et al., 2019), cocok digunakan untuk melakukan prediksi pada data non-linear dan data yang tidak stasioner (Abdoli et al., 2020). Metode LSTM ini diharapkan memperoleh hasil Prediksi kadar PM₁₀ yang baik sehingga berguna untuk mengetahui kondisi kualitas udara di hari-hari berikutnya dan berperan penting sebagai antisipasi pencemaran udara. Hasil prediksi kadar PM₁₀ di masa mendatang menjadi aspek kunci yang akan digunakan sebagai alat perbandingan antara kondisi RTH yang tersedia saat ini dengan RTH yang ideal, sehingga dapat kita tentukan kekurangan lahan RTH yang mampu mengurangi angka polutan di suatu wilayah.

METODE

Data Penelitian

Sumber data area studi pada penelitian ini adalah Kota Jambi yang merupakan ibu kota Provinsi Jambi. Data yang akan digunakan adalah data ISPU dari tanggal 1 Januari 2021 sampai dengan 30 Juni 2024. Data yang akan diolah adalah data harian angka PM₁₀ dari Stasiun Pemantauan Kualitas Udara (SPKU) Kota Jambi. Data ISPU bersumber dari Dinas Lingkungan Hidup Kota Jambi. Adapun atribut data ISPU yang digunakan pada penelitian ini adalah tanggal dan angka PM₁₀.

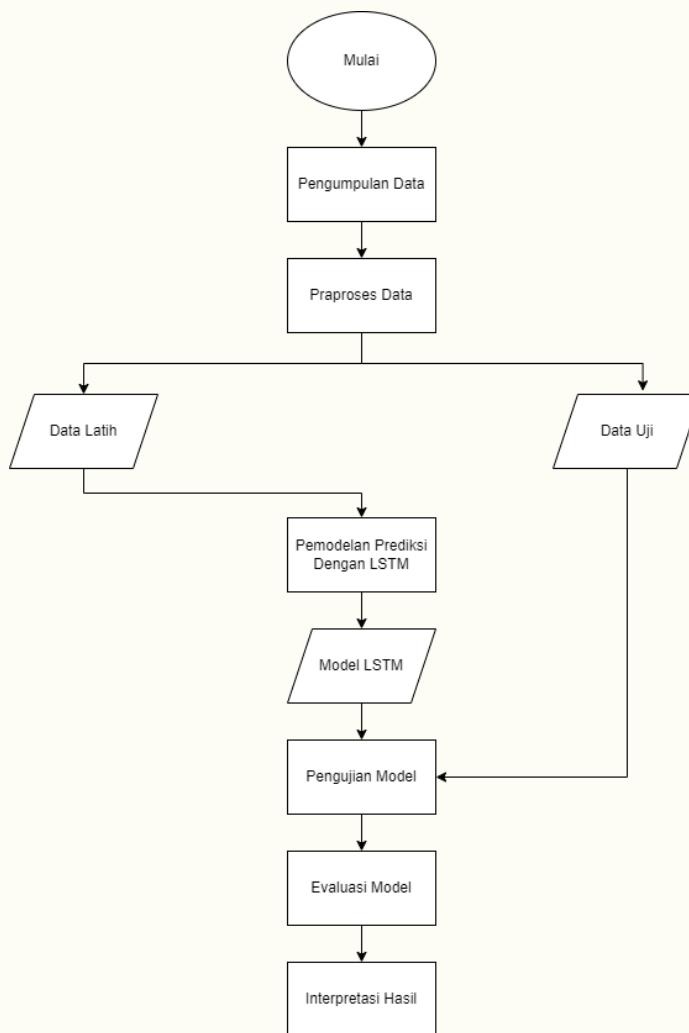
Alat Penelitian

Alat yang dibutuhkan dalam penelitian ini, yakni:

1. Perangkat keras (*Hardware*)
Laptop Asus dengan spesifikasi *processor Intel® Core™ i7-8550U CPU @ 1.780GHz, RAM 8 GB, dan system type 64-bit operating system*
2. Perangkat lunak (*Software*)
 - a. Sistem Operasi *Windows*
 - b. Jupyter notebook, Bahasa pemrograman *Python 3.7*
 - c. Library pandas, sklearn, keras, math, matplotlib, dan numpy untuk membuat model LSTM
 - d. Library streamlit untuk membuat GUI aplikasi web

Tahapan Penelitian

Tahapan penelitian disajikan pada Gambar 1. Penelitian ini terdiri dari beberapa tahap, yaitu pengumpulan data, praproses data, pembagian data, pembuatan model, dan evaluasi model.



Gambar 1. Tahapan Penelitian

Pengumpulan Data

Dalam melakukan penelitian ini, diperlukan pengumpulan dataset yang berupa kumpulan data *time series* kadar PM₁₀. Sampel data dari angka PM₁₀ dapat dilihat pada Tabel 1.

Tabel 1. Sampel Dataset Penelitian

| No. | Tanggal | Kadar PM ₁₀ ($\mu\text{g}/\text{m}^3$) |
|-----|---------------|---|
| 1. | 07 April 2024 | 18 |
| 2. | 08 April 2024 | 14 |
| 3. | 09 April 2024 | 16 |
| 4. | 10 April 2024 | 20 |
| 5. | 11 April 2024 | 21 |

Sumber : DLH Kota Jambi

Preprocessing Data

Pada tahapan ini dilakukan proses penghapusan data yang null, duplikasi data, dan penghapusan kolom yang tidak digunakan. Data pada penelitian ini memiliki rentang nilai yang berbeda untuk itu perlu dilakukan normalisasi data untuk menghindari nilai yang memiliki rentang nilai besar mendominasi nilai yang memiliki rentang nilai kecil. Pada penelitian ini, dilakukan normalisasi min-max dengan rentang 0-1(Kumar & Kumar, 2011). Rumus normalisasi min-max.

$$Norm = \frac{x - min(x)}{max(x) - min(x)} \dots 1$$

$min(x)$ = nilai minimum data

$max(x)$ = nilai maksimum data

Pembagian Data Latih dan Data Uji

Pada tahapan ini dilakukan proses pembagian dataset untuk dilakukan proses pelatihan dan data yang akan diuji. Data dibagi menjadi data latih dan data uji sebesar 80% data latih dan 20% data uji. Alokasi 80% untuk data latih umumnya memberikan volume data yang cukup signifikan bagi model LSTM untuk mempelajari pola, tren, dan musimanitas yang mendasari data historis konsentrasi PM10. Dengan set pelatihan yang lebih besar, model diharapkan dapat menangkap ketergantungan temporal yang kompleks dalam data deret waktu dengan lebih baik. Sementara itu, 20% data uji yang tidak terlihat sebelumnya berfungsi sebagai tolok ukur independen untuk mengevaluasi kemampuan generalisasi model, memberikan perkiraan yang lebih realistik tentang akurasi prediktifnya pada data baru. Pembagian 80/20 ini bertujuan untuk menyeimbangkan antara penyediaan data yang memadai untuk pembelajaran yang efektif dan mempertahankan ukuran set uji yang cukup untuk evaluasi kinerja yang andal. Data latih atau data *training* adalah kumpulan data yang memiliki label yang digunakan untuk mengenal karakteristik dari data tersebut oleh mesin agar dapat menghasilkan pola atau model dari data tersebut. Data latih digunakan untuk membuat permodelan, sedangkan data uji atau data *testing* adalah yang memiliki label untuk menguji performa dan kebenaran dari model yang telah dilatih (Shao et al., 2019).

Pemodelan dengan LSTM

Pemilihan Long Short-Term Memory (LSTM) untuk prediksi temporal kualitas udara, terutama konsentrasi PM₁₀, didasarkan pada kemampuannya yang unggul dalam memproses data berurutan dan menangkap ketergantungan jangka panjang, sebuah keunggulan yang tidak dimiliki oleh banyak algoritma machine learning lainnya. Data kualitas udara secara inheren bersifat temporal, di mana nilai saat ini dipengaruhi oleh nilai-nilai sebelumnya dan kondisi lingkungan. Meskipun Recurrent Neural Networks (RNN) dirancang untuk data sekuensial, RNN tradisional seringkali mengalami kesulitan dalam mempelajari ketergantungan jarak jauh akibat masalah vanishing/exploding gradient. LSTM mengatasi keterbatasan ini melalui arsitektur sel memori yang kompleks, yang memungkinkan mereka untuk menyimpan dan mengakses informasi dalam jangka waktu yang lama, menjadikannya lebih efektif dalam memodelkan pola temporal yang rumit dalam data PM₁₀.

LSTM memiliki *memory cell* dan *gate units* pada setiap *neurons* nya yang berfungsi untuk mengatur memori dalam setiap *neurons*. Terdapat empat proses fungsi aktivasi pada setiap masukan pada *neurons* yang selanjutnya disebut sebagai *gates units*. *Gates units* tersebut ialah *forget gates*, *input gates*, *cell gates*, dan *output gates*(Yu et al., 2019).

Dalam konstruksi LSTM, tiga jenis bobot digunakan. Bobot pada LSTM adalah bobot dari input layer ke hidden layer, bobot dari hidden layer ke output layer dan bobot dari contex layer ke hidden layer. Selain bobot, proses LSTM juga termasuk inisialisasi parameter pembelajaran, Pembelajaran dilakukan dengan menentukan bobot masing-masing node dalam jaringan. Parameter pembelajaran pada LSTM dalam penelitian ini adalah nilai learning rate, epoch dan activation function. Pembelajaran berhenti ketika nilai kesalahan mencapai target atau mencapai maksimum iterasi yang ditentukan.

Evaluasi

Setelah menyelesaikan proses pemodelan pada data, langkah selanjutnya adalah melakukan evaluasi terhadap kinerja model yang telah dibangun. Model prediksi PM₁₀ yang dibangun diuji menggunakan *Root Mean Square* (RMSE) dan *Mean Absolute Percentage Error* (MAPE). RMSE adalah metode untuk mengevaluasi teknik peramalan dengan mengukur tingkat akurasinya. Rumus dari RMSE yaitu:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad \dots 2$$

Keterangan :

- A_t : nilai data aktual
F_t : nilai hasil peramalan
n : banyak data

MAPE adalah rata-rata dari diferensiasi antara hasil sebenarnya dan nilai peramalan yang dinyatakan melalui persentase nilai aktual. Rumus dari MAPE yaitu :

$$PE_t = \left(\frac{e_t}{z_t} \right) \times 100 \quad \dots 3$$

Lalu,

$$MAPE = \frac{\sum_{t=1}^n |PE_t|}{n} \quad \dots 4$$

Keterangan :

- PE_t : persentase kesalahan pada waktu ke-t
E_t : nilai kesalahan waktu ke-t
Z_t : data aktual waktu ke-t
n : banyak data

Implementasi Graphic User Interface (GUI)

Setelah melakukan evaluasi pada model yang telah dikembangkan, langkah selanjutnya adalah tahap implementasi. Dalam membuat GUI prediksi angka PM₁₀, peneliti menggunakan framework streamlit. Streamlit merupakan sebuah framework berbasis python dan bersifat open-source yang dibuat untuk memudahkan dalam membangun aplikasi web di bidang sains data dan machine learning yang interaktif. GUI yang dibuat oleh peneliti untuk prediksi data indeks PM₁₀ dimana data yang di input yaitu data indeks PM₁₀ yang diambil dari DLH Kota Jambi dan output yang dihasilkan merupakan nilai hasil prediksi dan grafik hasil prediksi.

HASIL DAN PEMBAHASAN

Data Penelitian

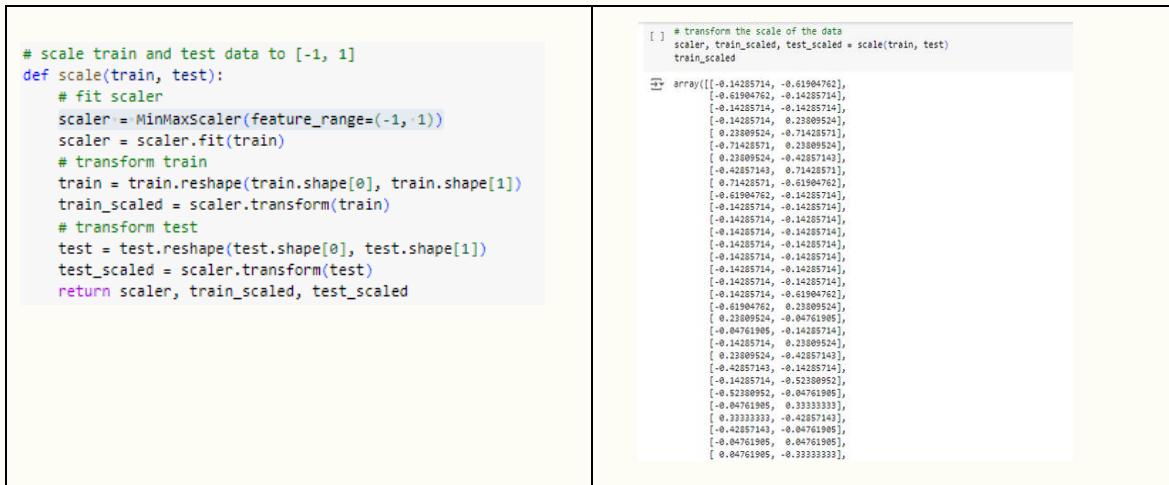
Area studi dan sumber data area studi pada penelitian ini adalah Kota Jambi. Data yang digunakan adalah data ISPU dari tanggal 1 Januari 2024 sampai dengan 30 Juni 2024. Data yang diolah adalah data harian indeks PM₁₀ dari Stasiun Pemantau Kualitas Udara (SPKU) di Kota Jambi. Data ISPU bersumber dari Dinas Lingkungan Hidup Kota Jambi berupa data dari 1 SPKU di Kota Jambi. Lokasi SPKU berada di lapangan kantor walikota Jambi. Peta lokasi stasiun pemantau dari Dinas Lingkungan Hidup Kota Jambi dapat diakses pada laman <https://ispu.menlhk.go.id/webv4/#/peta/JAMBI/-1.5818300247192383/103.5790023803711>.

Praproses Data

Tahapan praproses diawali dengan pembentukan dataset. Dataset terdiri dari PM₁₀, PM_{2,5}, SO₂, CO, O₃, NO₂, HC. Pada penelitian hanya berfokus pada data indeks PM₁₀. Data set ISPU dari Dinas Lingkungan Hidup Kota Jambi disajikan per jam, untuk menjadi data indeks perhari untuk perlu pengolahan data dengan mencari indeks PM₁₀ rata-rata perhari. Selanjutnya, tahapan praproses dilanjutkan dengan melakukan pemeriksaan dan penanganan missing value, dan normalisasi data(Alasadi & Bhaya, 2017). Pada penelitian ini, ditemukan sejumlah missing value pada data ISPU. Missing Value ini mengakibatkan ketidaklengkapan data dalam suatu model, sehingga menghambat analisis statistik yang akan dilakukan(Ren et al., 2023). Maka dari itu perlu metode untuk menangani permasalahan Missing Value tersebut dengan melakukan data imputation yang bertujuan untuk mengisi nilai-nilai pada data yang hilang dengan perkiraan atau estimasi yang sesuai seperti

menggunakan rata-rata, median, atau modus dari kolom data yang sesuai(Lin & Tsai, 2020). Pada penelitian data imputation menggunakan nilai rata-rata(Mohd Pauzi et al., 2021)

Setelah itu, data kemudian dinormalisasi menggunakan normalisasi min-max, dimana rentang nilai data diubah dari -1 sampai dengan 1(Henderi, 2021). Normalisasi data dilakukan untuk mencegah rentang nilai data yang terlalu jauh(Prasetyowati et al., 2022). Proses normalisasi data menggunakan Bahasa pemrograman python dan library sklearn.preprocessing. Data hasil praproses normalisasi dapat dilihat pada Gambar 2.



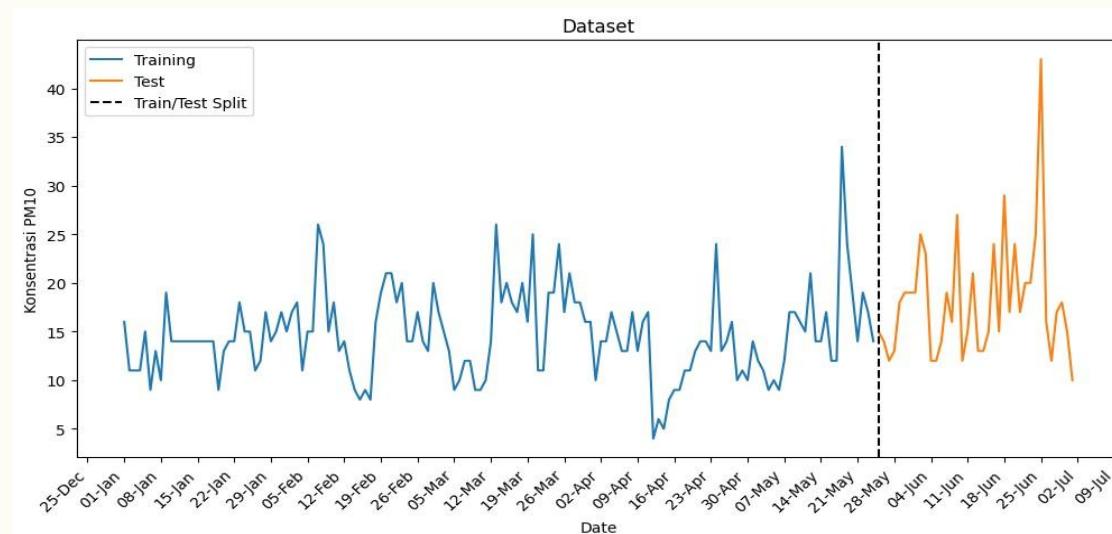
```
# scale train and test data to [-1, 1]
def scale(train, test):
    # fit scaler
    scaler = MinMaxScaler(feature_range=(-1, 1))
    scaler = scaler.fit(train)
    # transform train
    train = train.reshape(train.shape[0], train.shape[1])
    train_scaled = scaler.transform(train)
    # transform test
    test = test.reshape(test.shape[0], test.shape[1])
    test_scaled = scaler.transform(test)
    return scaler, train_scaled, test_scaled
```

```
[ ] # transform the scale of the data
scaler, train_scaled, test_scaled = scale(train, test)
train_scaled
```

```
array([[-0.14285714, -0.61904762],
       [-0.61904762, -0.14285714],
       [-0.14285714, -0.14285714],
       [-0.14285714, 0.23809524],
       [0.23809524, -0.71428571],
       [-0.71428571, 0.23809524],
       [0.23809524, -0.42857143],
       [-0.42857143, 0.71428571],
       [-0.71428571, -0.61904762],
       [-0.61904762, -0.14285714],
       [-0.14285714, -0.14285714],
       [-0.14285714, 0.14285714],
       [0.14285714, -0.14285714],
       [-0.14285714, -0.14285714],
       [-0.14285714, 0.14285714],
       [0.14285714, -0.14285714],
       [-0.14285714, -0.14285714],
       [-0.14285714, 0.14285714],
       [0.14285714, -0.14285714],
       [-0.14285714, -0.14285714],
       [-0.14285714, 0.23809524],
       [0.23809524, -0.42857143],
       [-0.42857143, 0.23809524],
       [-0.23809524, -0.52380952],
       [-0.52380952, -0.04761905],
       [-0.04761905, 0.33333333],
       [0.33333333, -0.42857143],
       [-0.42857143, -0.04761905],
       [-0.04761905, 0.64761905],
       [0.64761905, -0.33333333]]
```

Gambar 2. Proses Normalisasi Data

Data yang digunakan di dalam penelitian ini adalah data dari 1 SPKU di Kota Jambi dari tanggal 1 Januari 2024 sampai dengan 30 Juni 2024. Data ini kemudian dibagi menjadi data latih dan data uji dengan menggunakan proporsi 80% data latih dan 20% data uji. Data latih adalah data dari tanggal 1 Januari 2024 sampai dengan 23 Mei 2024, dan data uji adalah data tanggal 24 Mei 2024 sampai dengan 30 Juni 2024.

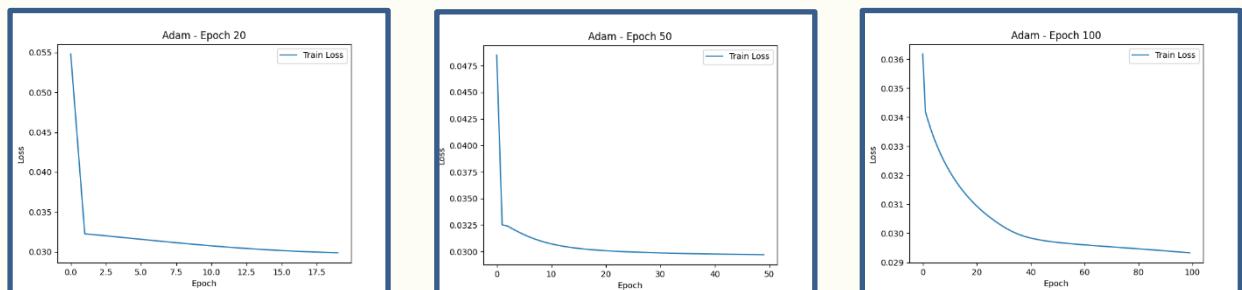


Gambar 3. Pembagian Data Latih dan Data Uji

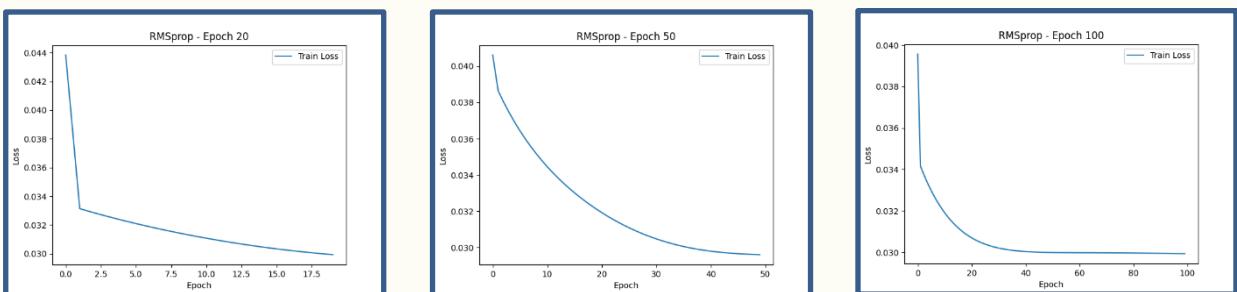
Pemodelan dengan LSTM

Model LSTM dibangun menggunakan library Keras yang ada pada bahasa pemrograman Python. Jumlah input node untuk model prediksi adalah sebanyak 1 node, 1 LSTM layer, 1 dropout layer dan 1 output layer, dimana nilai indeks polutan PM₁₀ diprediksi berdasarkan nilai indeks PM₁₀ di hari sebelumnya. Fungsi aktivasi yang digunakan untuk setiap layer LSTM adalah tanh, Keluaran dari

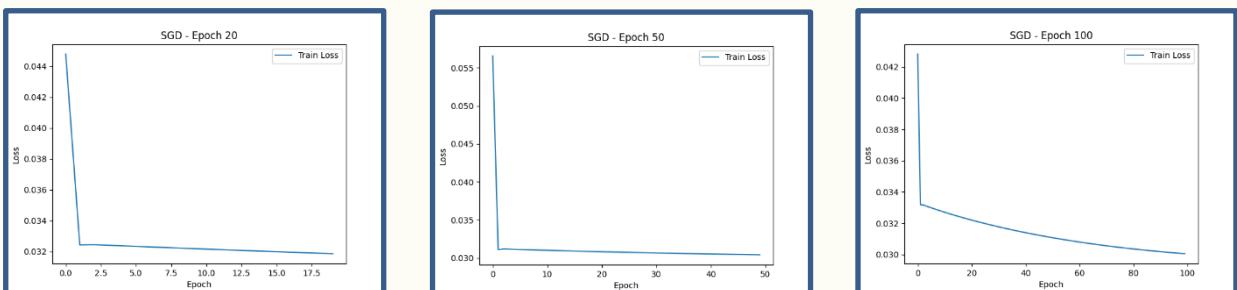
output layer adalah nilai indeks PM10 yang diprediksi. Adapun optimizer yang digunakan pada arsitektur pemodelan model adalah Adam, Adamax, RMSprop, dan SGD. Pada penelitian ini menggunakan epoch sebanyak 20, 50, dan 100 dan jumlah neuron sebanyak 7. Untuk mengetahui jumlah *epoch* yang tepat dapat dilihat nilai loss yang terkecil, dimana nilai loss merupakan nilai RMSE(Malsa et al., 2021). Berikut ini adalah hasil pengujian dari masing-masing optimizer yang dicobakan untuk epoch 20, 50, dan 100.



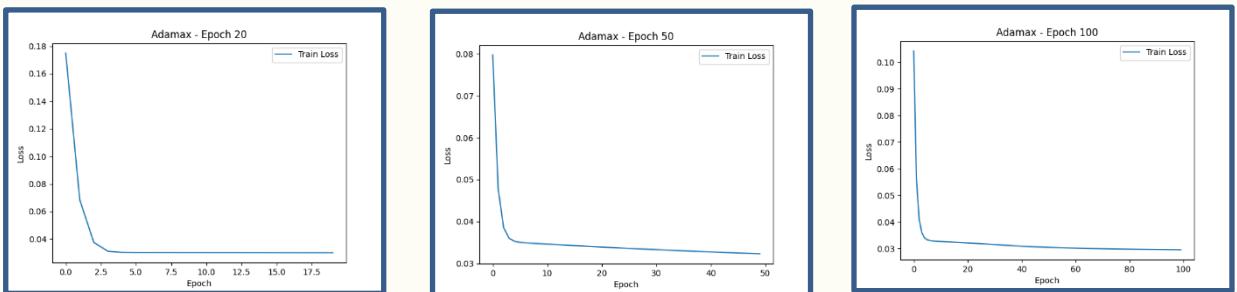
Gambar 4. Grafik Nilai Loss Optimizer Adam



Gambar 5. Grafik Nilai Loss Optimizer RMSprop



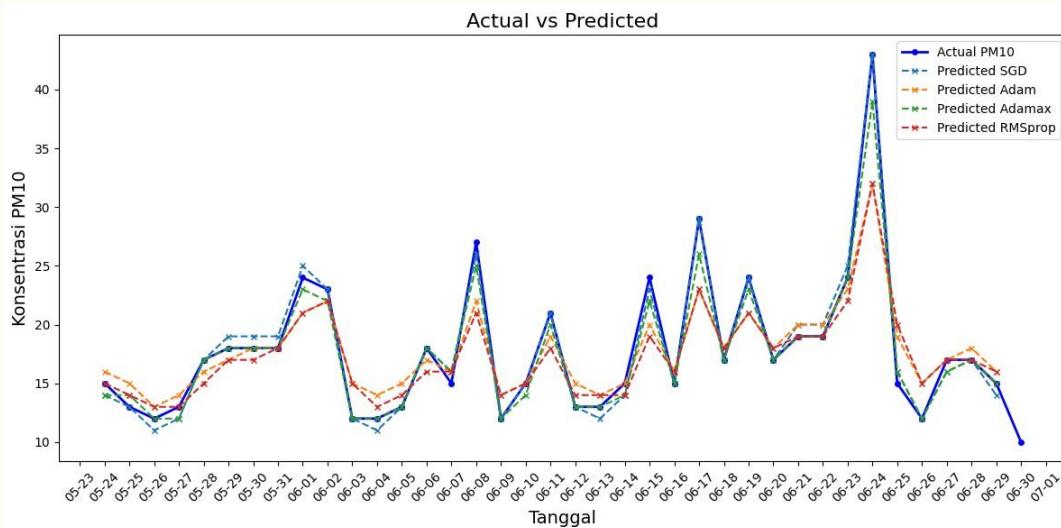
Gambar 6. Grafik Nilai Loss Optimizer SGD



Gambar 7. Grafik Nilai Loss Optimizer Adamax

Tabel 2. Perbandingan Hasil Evaluasi Optimizer

| Optimizer | Epoch | Loss | RMSE | MAPE |
|-----------------------------------|-------|--------|-------|--------|
| Adam | 20 | 0.0299 | 0.075 | 0.391% |
| | 50 | 0.0297 | 0.116 | 0.570% |
| | 100 | 0.0293 | 0.118 | 0.588% |
| RMSprop | 20 | 0.0299 | 0.060 | 0.329% |
| | 50 | 0.0296 | 0.092 | 0.487% |
| | 100 | 0.0299 | 0.097 | 0.497% |
| Stochastic Gradient Descent (SGD) | 20 | 0.0319 | 0.021 | 0.110% |
| Adamax | 50 | 0.0304 | 0.051 | 0.282% |
| | 100 | 0.0301 | 0.063 | 0.331% |
| | 20 | 0.0300 | 0.051 | 0.271% |
| RMSprop | 50 | 0.0330 | 0.022 | 0.124% |
| | 100 | 0.0285 | 0.099 | 0.541% |



Gambar 8. Hasil pengujian untuk masing-masing optimizer

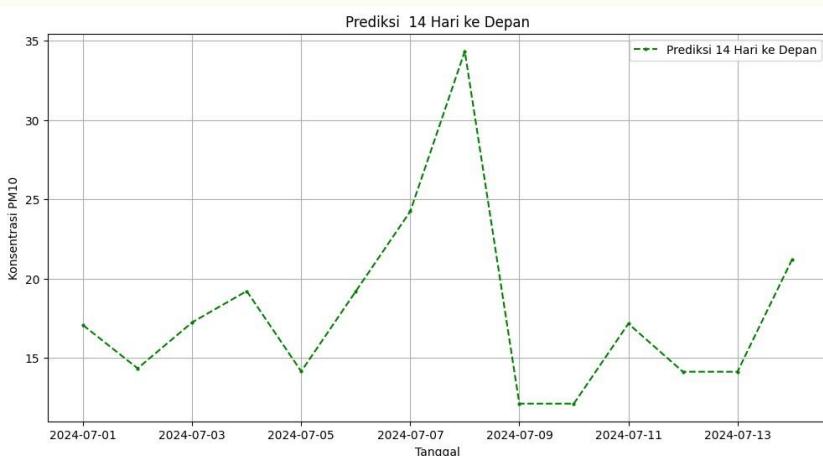
Hasil pengujian menunjukkan bahwa jumlah epoch (iterasi) 20 dengan optimizer SGD menghasilkan nilai RMSE yang terkecil yaitu 0,021 dan MAPE 0,11% jika dibandingkan dengan jumlah epoch dan optimizer lainnya sehingga model LSTM dengan kombinasi 20 epoch dan optimizer SGD merupakan model terbaik untuk kasus ini(Prasher et al., 2023). Hasil pengujian model LSTM dengan berbagai optimizer menunjukkan temuan yang menarik, di mana Stochastic Gradient Descent (SGD) menghasilkan kinerja prediksi konsentrasi PM₁₀ yang lebih unggul dibandingkan optimizer adaptif seperti Adam, Adamax, dan RMSprop. Meskipun SGD umumnya dikenal memiliki konvergensi yang lebih lambat dan sensitif terhadap pemilihan tingkat pembelajaran, dalam eksperimen ini, dengan konfigurasi tingkat pembelajaran dan momentum yang tepat, SGD mampu mencapai nilai RMSE dan MAPE yang lebih rendah. Keberhasilan SGD dalam kasus ini menyoroti pentingnya eksplorasi dan penyetelan hyperparameter yang teliti, bahkan pada algoritma optimasi klasik.

Implikasi dari hasil ini adalah bahwa tidak ada satu optimizer yang secara universal optimal untuk semua tugas. Pilihan optimizer sangat bergantung pada arsitektur model, karakteristik dataset, dan bahkan tugas prediksi yang spesifik. Meskipun optimizer adaptif sering menjadi pilihan pertama karena kemudahan penggunaan dan kinerja yang baik secara umum, eksplorasi dan penyetelan *hyperparameter* pada optimizer klasik seperti SGD masih dapat menghasilkan hasil yang kompetitif, bahkan unggul.

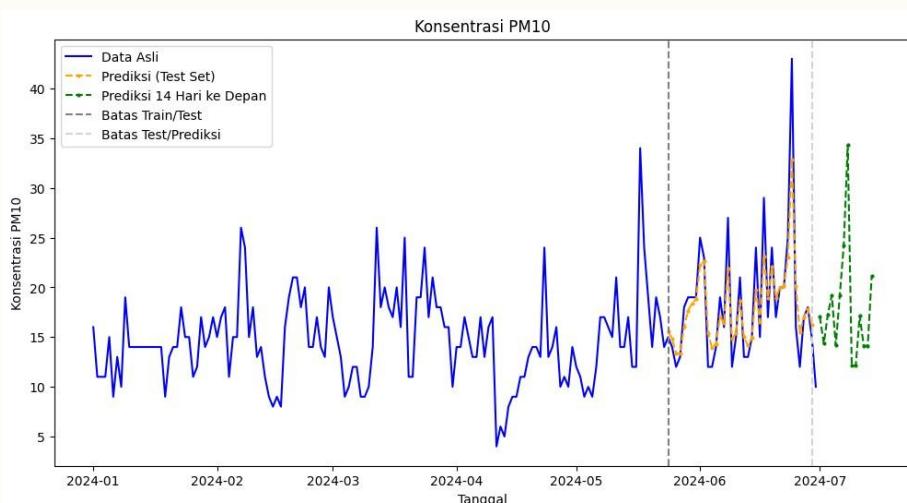
Prediksi Data

Model yang terbaik yaitu model LSTM dengan epoch 20 dan optimizer SGD digunakan untuk meramalkan nilai indeks PM₁₀ untuk 14 hari ke depan. Gambar 9 merupakan grafik visualisasi prediksi konsentrasi PM₁₀ untuk 14 hari ke depan. Gambar 10 merupakan visualisasi grafik perbandingan data aktual, prediksi, dan peramalan indeks PM₁₀. Tabel 3 menampilkan hasil prediksi konsentrasi PM₁₀ tanggal 1 hingga 14 Juli 2024. Berdasarkan peramalan konsentrasi PM₁₀ pada Tabel 3 tersebut diketahui bahwa kualitas udara di Kota Jambi untuk parameter PM₁₀ masih termasuk dalam kategori Baik karena range ISPU masih di bawah 50. Hal ini disebabkan di sepanjang tahun 2024 memang tidak ada kasus kebakaran hutan di Kota Jambi.

Berdasarkan UU No. 26 Tahun 2007 tentang Kawasan Hijau, luas RTH di Kota Jambi belum memenuhi luas RTH ideal yaitu 30% dari luas wilayah Kota Jambi. Adapun luas wilayah Kota Jambi yaitu 169,887 km² dan luas RTH yang tersedia di Kota Jambi yaitu 2.181,45 ha. Ketersediaan RTH hanya 12,84% dari luas wilayah Kota Jambi. Kurangnya ketersediaan RTH disebabkan oleh alih fungsi menjadi lahan terbangun, menjadi kawasan pemukiman, ruko, sarana dan prasarana infrastruktur kota



Gambar 9. Prediksi Indeks PM10 14 Hari Ke Depan



Gambar 10. Perbandingan Data Aktual, Prediksi, dan Peramalan indeks PM10

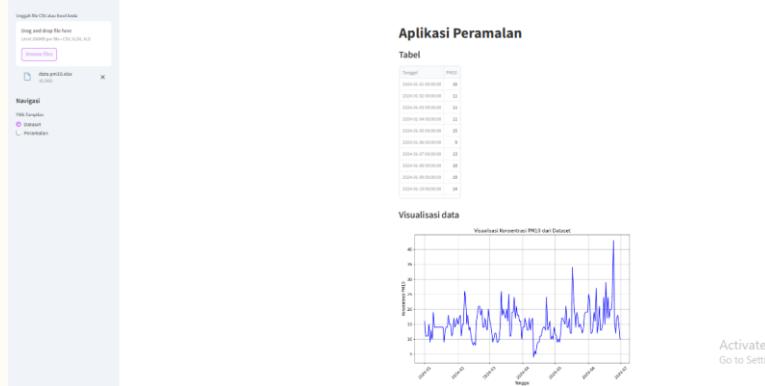
Tabel 3. Hasil Peramalan Indeks PM10

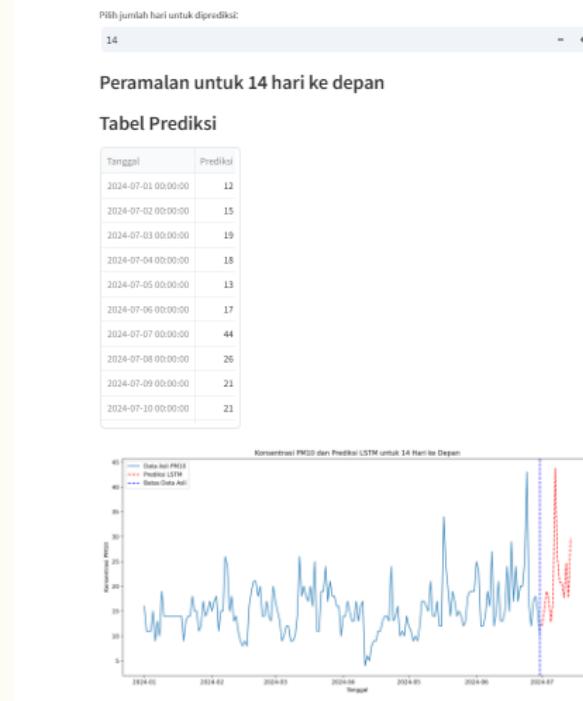
| Tanggal | Peramalan |
|------------|-----------|
| 01/07/2024 | 17 |
| 02/07/2024 | 14 |
| 03/07/2024 | 17 |
| 04/07/2024 | 19 |
| 05/07/2024 | 14 |
| 06/07/2024 | 19 |
| 07/07/2024 | 24 |
| 08/07/2024 | 34 |
| 09/07/2024 | 12 |
| 10/07/2024 | 12 |
| 11/07/2024 | 17 |
| 12/07/2024 | 14 |
| 13/07/2024 | 14 |
| 14/07/2024 | 21 |

Implementasi GUI (Graphic User Interface)

Pada penelitian ini, dirancang aplikasi peramalan yang dibangun melalui *library python Streamlit*. Aplikasi ini digunakan untuk meramalkan indeks PM10 kedepannya. Aplikasi ini hanya digunakan untuk melakukan sebuah proses *analisis time series*, dataset yang diolah hanya 1 variabel yakni dataset indeks PM10. Aplikasi dikembangkan agar pengguna yang tidak memahami statistika atau *non-programmer* dapat melakukan peramalan terhadap dataset indeks PM₁₀. Berikut kebutuhan sistem yang disediakan dalam aplikasi ini, yaitu Plot dataset *time series* dan Hasil peramalan LSTM.

Aplikasi ini dikembangkan dari beberapa *library* yang ada di *python*. Adapun *library-library* yang menunjang pengembangan aplikasi ini adalah *streamlit*, *numpy*, *pandas*, *matplotlib*, *pylab*, *sklearn*, *keras*. *Library-library* ini digunakan untuk metode atau model peramalan LSTM. Antarmuka yang pertama adalah pada bagian *sidebar* dan menu dataset. Pada menu ini pengguna dapat memasukkan dataset, melihat nama dataset, memilih menu dataset atau hasil peramalan, melihat tabel dataset yang mereka masukkan, variabel yang akan dipilih dan visualisasi dataset tersebut (Gambar 11). Antarmuka yang terakhir adalah menu hasil peramalan. Pada menu ini akan dilakukan proses peramalan dataset yang telah dimasukkan sebelumnya. Pada menu ini akan ditampilkan hasil prediksi berupa visualisasi data perbandingan data uji dan data tes dari LSTM, selain itu, ditampilkan juga nominal angka perbandingannya agar memudahkan pengguna. Kemudian, ringkasan dari model terbaik serta peramalan kedepannya (Gambar 12).





Gambar 12. Menu Hasil Peramalan

KESIMPULAN

Hasil pengujian model LSTM dengan berbagai optimizer menunjukkan temuan yang menarik, di mana Stochastic Gradient Descent (SGD) dengan 20 epoch menghasilkan kinerja prediksi konsentrasi PM₁₀ yang lebih unggul dibandingkan optimizer adaptif seperti Adam, Adamax, dan RMSprop. Nilai RMSE yang terkecil yaitu 0,021 dan MAPE 0,11% yang berarti bahwa peramalan sangat akurat. Berdasarkan peramalan konsentrasi PM₁₀ diketahui bahwa kualitas udara di Kota Jambi untuk parameter PM₁₀ masih termasuk dalam kategori Baik karena range PM₁₀ masih di bawah 50.

DAFTAR PUSTAKA

- Abdoli, G., MehrAra, M., & Ebrahim Ardalani, M. (2020). Comparing The Prediction Accuracy Of Lstm And Arima Models For Time-Series With Permanent Fluctuation. *Gênero & Direito*, 9(2). <https://doi.org/10.22478/ufpb.2179-7137.2020v9n2.50782>
- Aji Tritamtama, K., Sembiring, F. E. S., Choiruddin, A., & Patria, H. (2023). Analysis of Air Pollution (SO₂) at Some Point of Congestion in DKI Jakarta. *Disease Prevention and Public Health Journal*, 17(1), 82–92. <https://doi.org/10.12928/dpphj.v17i1.6147>
- Alasadi, S. A., & Bhaya, W. S. (2017). Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences*. <https://doi.org/10.3923/jeasci.2017.4102.4107>
- Chiam, Z., Song, X. P., Lai, H. R., & Tan, H. T. W. (2019). Particulate matter mitigation via plants: Understanding complex relationships with leaf traits. *Science of The Total Environment*, 688, 398–408. <https://doi.org/10.1016/j.scitotenv.2019.06.263>
- Gupta, A., Moniruzzaman, M., Hande, A., Rousta, I., Olafsson, H., & Mondal, K. K. (2020). Estimation of particulate matter (PM_{2.5}, PM₁₀) concentration and its variation over urban sites in Bangladesh. *SN Applied Sciences*. <https://doi.org/10.1007/s42452-020-03829-1>
- Henderi, H. (2021). Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. *IJIIS: International Journal of Informatics and Information Systems*, 4(1), 13–20. <https://doi.org/10.47738/ijiis.v4i1.73>

- Huang, R., Wei, C., Wang, B., Yang, J., Xu, X., Wu, S., & Huang, S. (2022). Well performance prediction based on Long Short-Term Memory (LSTM) neural network. *Journal of Petroleum Science and Engineering*, 208, 109686. <https://doi.org/10.1016/j.petrol.2021.109686>
- Joshi, N., Joshi, A., & Bist, B. (2020). Phytomonitoring and Mitigation of Air Pollution by Plants. In *Sustainable Agriculture in the Era of Climate Change* (pp. 113–142). Springer International Publishing. https://doi.org/10.1007/978-3-030-45669-6_5
- Kumar, J. Y., & Kumar, B. S. (2011). Min max normalization based data perturbation method for privacy protection. *International Journal of Computer & Communication Technology*.
- Latif, N., Selvam, J. D., Kapse, M., Sharma, V., & Mahajan, V. (2023). Comparative Performance of LSTM and ARIMA for the Short-Term Prediction of Bitcoin Prices. *Australasian Accounting, Business and Finance Journal*, 17(1), 256–276. <https://doi.org/10.14453/aabfj.v17i1.15>
- Lin, W.-C., & Tsai, C.-F. (2020). Missing value imputation: a review and analysis of the literature (2006–2017). *Artificial Intelligence Review*, 53(2), 1487–1509. <https://doi.org/10.1007/s10462-019-09709-4>
- Maduna, K., & Tomašić, V. (2017). Air pollution engineering. *Physical Sciences Reviews*, 2(12). <https://doi.org/10.1515/psr-2016-0122>
- Malsa, N., Vyas, V., & Gautam, J. (2021). RMSE calculation of LSTM models for predicting prices of different cryptocurrencies. *International Journal of System Assurance Engineering and Management*. <https://doi.org/10.1007/s13198-021-01431-1>
- Mangones, S. C., Jaramillo, P., Rojas, N. Y., & Fischbeck, P. (2020). Air pollution emission effects of changes in transport supply: the case of Bogotá, Colombia. *Environmental Science and Pollution Research*, 27(29), 35971–35978. <https://doi.org/10.1007/s11356-020-08481-1>
- Meo, S. A., Almutairi, F. J., Abukhalaf, A. A., & Usmani, A. M. (2021). Effect of Green Space Environment on Air Pollutants PM2.5, PM10, CO, O₃, and Incidence and Mortality of SARS-CoV-2 in Highly Green and Less-Green Countries. *International Journal of Environmental Research and Public Health*, 18(24), 13151. <https://doi.org/10.3390/ijerph182413151>
- Mohd Pauzi, N. A., Wah, Y. B., Deni, S. M., Abdul Rahim, S. K. N., & Suhartono. (2021). Comparison of Single and MICE Imputation Methods for Missing Values: A Simulation Study. *Pertanika Journal of Science and Technology*, 29(2). <https://doi.org/10.47836/pjst.29.2.15>
- Prasetyowati, S. A. D., Ismail, M., Budisusila, E. N., Setiadi, D. R. I. M., & Purnomo, M. H. (2022). Dataset Feasibility Analysis Method based on Enhanced Adaptive LMS method with Min-max Normalization and Fuzzy Intuitive Sets. *International Journal on Electrical Engineering and Informatics*, 14(1), 55–75. <https://doi.org/10.15676/ijeei.2022.14.1.4>
- Prasher, S., Nelson, L., & Jagdish, M. (2023). Potato leaf disease prediction using RMSProp, Adam and SGD optimizers. *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, 343–347. <https://doi.org/10.1109/InCACCT57535.2023.10141714>
- Priyankara, S., Senarathna, M., Jayaratne, R., Morawska, L., Abeysundara, S., Weerasooriya, R., Knibbs, L. D., Dharmage, S. C., Yasaratne, D., & Bowatte, G. (2021). Ambient PM2.5 and PM10 Exposure and Respiratory Disease Hospitalization in Kandy, Sri Lanka. *International Journal of Environmental Research and Public Health*, 18(18). <https://doi.org/10.3390/ijerph18189617>
- Ren, L., Wang, T., Sekhari Seklouli, A., Zhang, H., & Bouras, A. (2023). A review on missing values for main challenges and methods. In *Information Systems*. <https://doi.org/10.1016/j.is.2023.102268>
- Saputra, R., Handika, R. A., & Lestari, R. A. (2019). Analisis Sebaran Polutan Particulate Matter (PM10) Pada Harian Musim Kemarau di Kota Jambi. *Jurnal Engineering*, 1(1). <https://doi.org/10.22437/jurnalengineering.v1i1.6276>

- Shao, S., McAleer, S., Yan, R., & Baldi, P. (2019). Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Transactions on Industrial Informatics*. <https://doi.org/10.1109/TII.2018.2864759>
- Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). A Comparison of ARIMA and LSTM in Forecasting Time Series. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*. <https://doi.org/10.1109/ICMLA.2018.00227>
- Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L., & Cheng, Z. (2020). Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *Journal of Petroleum Science and Engineering*, 186, 106682. <https://doi.org/10.1016/j.petrol.2019.106682>
- Vlasov, D., Ramírez, O., & Luhar, A. (2022). Road Dust in Urban and Industrial Environments: Sources, Pollutants, Impacts, and Management. In *Atmosphere*. <https://doi.org/10.3390/atmos13040607>
- Wang, Y., Yuan, Y., Wang, Q., Liu, C. G., Zhi, Q., & Cao, J. (2020). Changes in air quality related to the control of coronavirus in China: Implications for traffic and industrial emissions. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.139133>
- Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, 31(7), 1235–1270. https://doi.org/10.1162/neco_a_01199