

Optimasi Long Short Term Memory Dengan Adam Menggunakan Data Udara Kota DKI Jakarta

Arvi Arkadia¹, Bayu Hananto², Desta Sandya Prasvita³
Prodi S1 Informatika / Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jakarta

Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Daerah Khusus Ibukota Jakarta 12450
rvarkdy@gmail.com¹, bayuhananto@upnvj.ac.id², desta.sandya@upnvj.ac.id³

Abstrak. Udara merupakan elemen penting bagi kehidupan. Pertumbuhan ekonomi dan urbanisasi yang terjadi di daerah kota besar berpotensi meningkatkan penggunaan energi sehingga mengakibatkan pencemaran udara. Di Indonesia memutuskan lima parameter yang dipergunakan sebagai zat polutan yang mengakibatkan pencemaran udara, zat tersebut diantaranya Nitrogen Dioksida (NO₂), Sulfur Dioksida (SO₂), Partikel Debu (PM₁₀), Ozon (O₃), dan Karbon Monoksida (CO). Kelima zat tersebut menjadi tolak ukur pada penentuan taraf udara yang ada di Indeks Standar Pencemaran Udara (ISPU). Metode Long Short Term Memory digunakan pada penelitian ini sebagai model yang digunakan dalam pengolahan data time series. Model LSTM digunakan untuk prediksi mengenai kualitas udara dengan kesalahan komputasi yang minimum. Penggunaan model LSTM dengan menggunakan Adam Optimizer untuk pengoptimalan nilai setiap layer agar menghasilkan prediksi yang akurat. Hasil prediksi dengan tingkat akurasi MAPE pada parameter PM₁₀ sebesar 4,37%, parameter SO₂ sebesar 5,02%, parameter CO sebesar 18,50%, parameter O₃ sebesar 5,23%, dan parameter NO₂ sebesar 37,28%.

Kata Kunci: LSTM, ISPU, dan Adam

1 Pendahuluan

Udara adalah unsur utama bagi kehidupan makhluk yang ada di muka bumi ini. Udara mengandung oksigen yang sangat dibutuhkan oleh tubuh. Pertumbuhan ekonomi dan urbanisasi yang terjadi di daerah kota besar berpotensi meningkatkan penggunaan energi listrik, air, dan minyak bumi. Hal tersebut mengakibatkan pencemaran udara yang membuat buruk kualitas udara di daerah kota besar seperti Jakarta. Polusi udara menciptakan penurunan kualitas udara yang sangat berpengaruh bagi manusia dan lingkungannya serta masalah Kesehatan yang ditimbulkan dari pencemaran udara [13]. Di Indonesia menetapkan lima parameter yang digunakan sebagai zat polutan yang mengakibatkan pencemaran udara, zat polutan tersebut diantaranya Nitrogen Dioksida (NO₂), Sulfur Dioksida (SO₂), Partikel Debu (PM₁₀), Ozon (O₃), dan Karbon Monoksida (CO). Kelima zat polutan ini menjadi tolak ukur dalam menentukan taraf udara pada Indeks Standar Pencemaran Udara (ISPU). ISPU yaitu kadar yang tidak memiliki ukuran atau dimensi yang bermanfaat dalam merepresentasikan ambien taraf udara di daerah dalam beberapa waktu yang berdampak bagi kebugaran makhluk hidup. Pada setiap parameter ISPU memiliki standar nilai pada rentang tertentu, dimana kategori baik pada 0 µg/m³ hingga 50 µg/m³, sedang pada 51 µg/m³ hingga 100 µg/m³, tidak sehat pada 101 µg/m³ hingga 200 µg/m³, dan sangat tidak sehat memiliki nilai lebih dari 200 µg/m³. Kategori kesehatan udara pada ISPU sendiri terdiri dari salah satu parameter yang memiliki ambang batas nilai tertinggi yang akan dikategorikan sebagai penentu kualitas udara.

Perkembangan dalam dunia *Artificial Intelligence* semakin pesat dan terus diaplikasikan dalam penyelesaian permasalahan yang terjadi di kehidupan sehari-hari. Dengan membangun model dalam memprediksi parameter ISPU di DKI Jakarta dapat berguna bagi pemangku kepentingan yang ada dalam bidang lingkungan hidup yang menangani permasalahan lingkungan hidup. *Long Short Term Memory* (LSTM) digunakan dalam pembuatan model prediksi dengan *data time series* kualitas udara untuk menghasilkan akurasi yang tinggi dan hasil evaluasi pengujian yang rendah. Penggunaan LSTM cocok untuk memproses, membuat prediksi, dan mengklasifikasi dalam *data time series*, karena ada kemungkinan durasi yang tidak dapat diketahui diantara kejadian dalam data tersebut yang penting dalam rangkaian waktu tertentu [1].

Dari penelitian sebelumnya menyatakan data mengenai kualitas udara dalam data runtun waktu (*stochastic time series*) sehingga pemrosesan prediksi didasarkan pada data rangkaian waktu atau historis dalam kualitas udara.

Analisis *time series* merupakan proses statistik yang digunakan dalam memproses prediksi struktur dari probabilitas kejadian yang akan datang yang dimanfaatkan untuk pengambilan keputusan. *Recurrent Neural Network* (RNN) menangani permasalahan dalam urutan data *time series* karena memiliki jaringan yang membentuk siklus terarah dengan maksud untuk mempertahankan suatu nilai dari satu iterasi ke iterasi berikutnya dengan menggunakan output sebagai nilai input untuk proses berikutnya. Jaringan sederhana dari RNN bekerja dengan baik dalam kasus memori jangka pendek tapi memiliki masalah dalam proses memori jangka panjang yang membuat *gradient* menghilang. Sehingga *Long Short Term Memory* (LSTM) dikembangkan untuk mengatasi permasalahan *gradient* yang hilang [24]. Algoritma *Machine Learning* dan *Deep Learning* banyak digunakan dalam memprediksi kualitas udara. Prediksi dari salah satu gas polutan PM_{2.5} untuk mengetahui tingkat konsentrasinya menggunakan metode *Long Short Term Memory* (LSTM), *Linear Regression* (LR), dan *Artificial Neural Network* (ANN). Dimana dalam penelitian tersebut performa LSTM lebih baik dalam memprediksi gas polutan PM_{2.5} [12]. Model LSTM digunakan dalam pembuatan prediksi mengenai *Air Quality Index* dengan kesalahan komputasi yang minimum [7]. Penggunaan model LSTM dengan menggunakan *Adam Optimizer* untuk pengoptimalan bobot setiap layer agar menghasilkan prediksi yang akurat. Untuk mengoptimalkan bobot setiap layer diperlukan *Adam Optimizer* yang menunjukkan kecepatan konvergen daripada menggunakan *Stochastic Gradient Descent* (SGD). *Gradient Descent* dimodifikasi dengan Algoritma *Adam Optimizer* yang mana secara komputasi efisien dan sesuai untuk optimasi dengan parameter yang besar. Output layer dari Adam memberikan prediksi yang lebih akurat [5].

Pada penelitian Elampartihi, et al (2021) dengan penelitian pendekatan deep learning untuk prediksi *air quality index* dengan PM_{2.5}. Dalam penelitian tersebut menggunakan data dari kota Chennai dengan metode *Long Short Term Memory*. Dari penelitian tersebut diperoleh nilai R² pada training dan testing sebesar 0,632 dan 0,570 yang mana mengindikasikan model LSTM cocok digunakan dalam prediksi PM_{2.5}. Berdasarkan uraian penelitian yang telah dijelaskan memiliki akurasi yang sangat baik, penggunaan metode *Long Short Term Memory* (LSTM) dengan *Adam Optimizer* digunakan pada penelitian ini dengan maksud untuk memprediksi kualitas udara pada ISPU dengan lima parameter yang digunakan. Dengan uraian masalah penggunaan *Adam Optimizer* dalam mengoptimasi model LSTM pada keakuratan nilai akurasi yang dihasilkan oleh model LSTM serta perbandingan hasil model LSTM dengan varian dari *Adam Optimizer*. Sehingga penelitian ini memiliki tujuan dalam pembuatan model LSTM dengan *Adam Optimizer* dan perhitungan nilai akurasi dari penggunaan model LSTM dengan Adam.

2 Tinjauan Pustaka

2.1 Pencemaran Udara

Udara adalah sekumpulan gas dimana komposisi dengan jumlah besar dalam udara yaitu oksigen dan nitrogen. Oksigen memiliki peranan penting dalam menunjang kehidupan makhluk hidup dan reaksi kimia dalam zat lain. Komposisi gas yang ada dalam udara bermacam-macam dalam satu wilayah tertentu. Polusi Udara merupakan sekumpulan partikulat asing yang terkandung pada komposisi udara yang mampu mengubah komposisi udara menjadi tidak normal. Adanya partikulat tersebut dengan volume tertentu dalam jangka waktu lama dapat mengakibatkan kerusakan ekosistem makhluk hidup [21]. Menurut Prabowo dan Muslim menyatakan bahwa polusi udara merupakan proses menyusupnya komponen lain, energi atau partikel pada udara karena pengaruh aktivitas manusia yang mengakibatkan kesehatan udara mengalami penurunan dengan tingkat kesehatan yang berbeda. Pencemaran udara dapat terjadi karena faktor pendukung yang menyebabkan proses pencemaran. Faktor pendukung terciptanya proses polusi udara adanya zat polutan yang mengandung gas polutan, proses interaksi bahan pencemaran dalam udara menyebabkan penurunan kualitas udara dan efek negatif pada lingkungan. Zat pencemar dalam udara dapat digolongkan menjadi gas atau partikulat. Dalam bentuk gas digolongkan menjadi : (1) golongan gas berbahaya (uap air raksa, Vinil Klorida, Benzene), (2) golongan beracun (Hidrogen Sulfida, Sulfur Dioksida), (3) golongan nitrogen (Nitrogen Dioksida, Amoniak, Nitrogen Monoksida, Karbon Dioksida), (4) golongan karbon (Hidrokarbon) [13].

2.2 Indeks Standar Pencemaran Udara (ISPU)

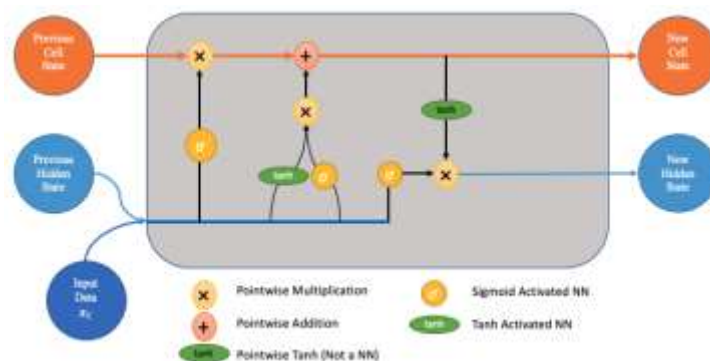
Indeks Standar Pencemaran Udara (ISPU) ialah hasil proses perhitungan pada kualitas udara yang berguna bagi masyarakat dalam menjelaskan tingkat kesehatan udara dan dampak bagi kesehatan manusia apabila udara tersebut dihirup dalam kurun waktu tertentu. Penentuan ISPU berdasarkan tingkat mutu udara pada hewan, tumbuhan, bangunan, dan manusia. Pada Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Pasal 2 Ayat 2 Nomor 14 Tahun 2020, dalam ISPU terdapat lima gas polutan yang dikategorikan sebagai bahan yang mengkontaminasi udara, yaitu : Sulfur Dioksida (SO₂), Partikel Debu (PM₁₀), Karbon Monoksida (CO), Nitrogen Dioksida (NO₂), dan Ozon (O₃). Indeks Standar Pencemaran Udara ditetapkan berdasarkan pengubahan konsentrasi bahan pencemar yang diperhitungkan dalam suatu angka yang tidak memiliki dimensi. Rentang nilai ISPU dan kategori dipaparkan pada Tabel 1.

Tabel 1. Kategori ISPU

Kategori	Rentang Nilai	Keterangan
Baik	0 - 50	Tidak memberikan efek kepada manusia atau makhluk lain.
Sedang	51 – 100	Tidak terlalu berpengaruh bagi manusia tetapi berpengaruh bagi tanaman yang sensitif.
Tidak Sehat	101 – 199	Bersifat merugikan bagi manusia, hewan dan tumbuhan.
Sangat Tidak Sehat	200 – 299	Berpengaruh pada kesehatan makhluk hidup dalam populasi yang terdampak.
Berbahaya	>300	Sangat berpengaruh bagi kesehatan dan menimbulkan masalah kesehatan serius secara umum.

2.3 Long Short Term Memory (LSTM)

Long Short Term Memory merupakan jaringan yang dibangun secara khusus dalam menangani permasalahan jangka panjang. Jaringan LSTM banyak digunakan dalam data terurut yang kompleks dan bidang analisis pola seperti pemodelan akustik skala besar dan pemodelan bahasa [7]. Long Short Term Memory menunjukkan sebuah informasi tetap ada dalam jaringan berurutan. Jaringan LSTM dirancang secara khusus dalam mengatasi permasalahan ketergantungan jangka panjang yang ada dalam RNN [24]. Perbedaan dasar dalam membedakan struktur RNN dan jaringan LSTM adalah pada modul yang berulang dan jaringan LSTM menggunakan struktur gate dalam menambah atau menghapus informasi dalam status sel [20]. LSTM mampu memecahkan masalah terhadap gradient descent dalam metode RNN. LSTM memiliki sebuah koneksi yang memberikan umpan balik sehingga membuatnya berbeda dengan jaringan feedforward neural network. Fungsi ini memungkinkan LSTM dalam memproses keseluruhan data yang telah terurut tanpa mengidentifikasi setiap elemen dalam urutan data secara independen, tetapi mempertahankan informasi penting pada data sebelumnya untuk digunakan dalam pemrosesan elemen data baru [27]. LSTM memiliki kinerja lebih mudah dalam mengolah data lampau dalam suatu memori. Pada umumnya, LSTM digunakan sebagai metode untuk memprediksi, memproses, dan mengklasifikasikan mengenai data time series dalam kurun waktu tertentu atau tidak diketahui dalam memori data [10]. Struktur LSTM digambarkan dalam Gambar 1.



Gambar 1. Struktur LSTM

Secara operasi matematis yang dilakukan pada setiap lapisan dalam jaringan LSTM pada Gambar 1 dijelaskan dalam sebagai berikut:

1. lapisan *sigmoid* paling kiri disebut sebagai *forget gate* yang mana bertugas dalam memutuskan informasi apa mengenai status sel yang perlu disimpan.

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \dots\dots\dots (1)$$

Dimana nilai f merupakan output dari jaringan *sigmoid*, h merupakan *output* dari lapisan sebelumnya, x merupakan nilai *input* dan b merupakan nilai *bias*.

2. Terdapat dua langkah proses dalam pemutusan informasi apa yang harus ditambahkan dalam status sel. Hal tersebut dilakukan oleh lapisan *sigmoid* lain yang dikenal dengan *input gate* dan lapisan *tanh*.

$$i_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \dots\dots\dots (2)$$

Dimana nilai i merupakan *output* dari *input gate*. Setelah melewati perhitungan dari *input gate* yang bertugas dalam menghitung nilai-nilai yang harus diperbaharui dari lapisan *tanh* untuk menciptakan nilai vektor baru yang akan dimasukan kedalam status sel.

$$\tilde{C}_t = \tanh (W_C \cdot [h_{t-1}, x_t] + b_C) \dots\dots\dots (3)$$

Dimana nilai C mengacu pada status sel.

3. Pembaharuan terhadap status sel dengan melupakan sejumlah informasi yang telah diperhitungkan dari penggunaan *forget gate* yaitu f_t dan penambahan nilai baru yaitu \tilde{C}_t dengan jumlah yang terhitung dari penggunaan *input gate* yaitu i_t .

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \dots\dots\dots (4)$$

4. Proses akhir dengan menghitung output dari filterisasi status sel dan hanya menyimpan informasi yang dibutuhkan dan relevan. Hal tersebut dijalankan dalam dua proses. Pertama, penggunaan lapisan *sigmoid* yang dikenal dengan *output gate* untuk memutuskan bagian mana dari status sel yang menghasilkan sebuah keluaran dan dikalikan dengan vektor dengan status sel yang telah melewati lapisan *tanh*.

$$O_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \dots\dots\dots (5)$$

Dimana nilai O_t merupakan *output* dari *output gate*.

$$h_t = o_t * \tanh(C_t) \dots\dots\dots (6)$$

2.4 Adaptive Moment Estimation (Adam)

Pada tahun 1940-an, matematik programming identik dengan penggunaan optimasi. Permasalahan mengenai optimasi merupakan tujuan untuk memaksimalkan atau meminimalkan dengan pemilihan nilai input dari sekumpulan nilai yang ada [25]. Saat ini, optimasi merupakan istilah yang sering digunakan pada lingkup AI. Khususnya permasalahan yang terjadi pada *deep learning*. Salah satu optimasi yang ada pada *deep learning* adalah *Adaptive Moment Estimation* (Adam). Algoritma Adam pertama kali dipublikasikan oleh Kingma dalam ICLR pada tahun 2015 dari paper dengan judul “*Adam: A Method for Stochastic Optimization*”, adam didefinisikan sebagai sebuah metode yang mampu melakukan optimasi pada stokastik dengan meminimumkan bobot nilai yang hanya membutuhkan gradien orde pertama dalam memori yang kecil. Secara struktural, Adam merupakan kombinasi dari RMSprop dan *Stochastic Gradient Descent* (SGD) dengan momentum. Adam menggunakan gradien kuadrat dalam melakukan skala tingkat pembelajaran seperti RMSprop dan memanfaatkan momentum dari rata-rata pergerakan gradien seperti SGD dengan momentum. Adam merupakan metode learning rate adaptif dimana penghitungan learning rate untuk individu memiliki parameter yang berbeda. Nama tersebut berasal dari adaptive moment estimation dengan alasan bahwa adam menggunakan estimasi momen pada gradien pertama dan gradien kedua dalam menyesuaikan learning rate pada setiap bobot dari neural network [3].

```

input :  $\gamma$  (lr),  $\beta_1, \beta_2$  (betas),  $\theta_0$  (params),  $f(\theta)$  (objective)
          $\lambda$  (weight decay), amsgrad, maximize
initialize :  $m_0 \leftarrow 0$  (first moment),  $v_0 \leftarrow 0$  (second moment),  $\widehat{v}_0^{max} \leftarrow 0$ 

```

```

for  $t = 1$  to ... do
  if maximize :
     $g_t \leftarrow -\nabla_{\theta} f_t(\theta_{t-1})$ 
  else
     $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$ 
  if  $\lambda \neq 0$ 
     $g_t \leftarrow g_t + \lambda \theta_{t-1}$ 
   $m_t \leftarrow \beta_1 m_{t-1} + (1 - \beta_1) g_t$ 
   $v_t \leftarrow \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$ 
   $\widehat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ 
   $\widehat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ 
  if amsgrad
     $\widehat{v}_t^{max} \leftarrow \max(\widehat{v}_t^{max}, \widehat{v}_t)$ 
     $\theta_t \leftarrow \theta_{t-1} - \gamma \widehat{m}_t / (\sqrt{\widehat{v}_t^{max}} + \epsilon)$ 
  else
     $\theta_t \leftarrow \theta_{t-1} - \gamma \widehat{m}_t / (\sqrt{\widehat{v}_t} + \epsilon)$ 

```

```

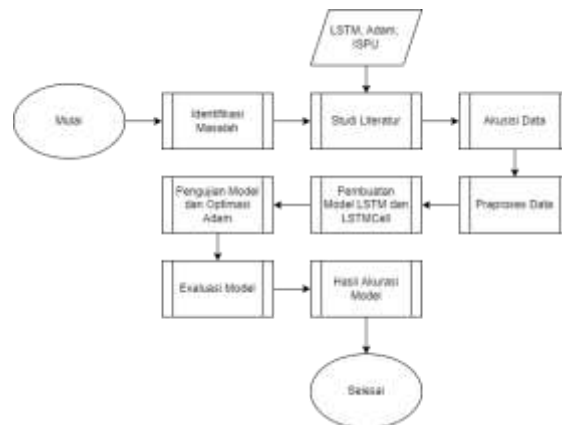
return  $\theta_t$ 

```

Gambar 2. Algoritma Adam

3 Metodologi

Dalam penyusunan serta penulisan pada pengkajian jurnal ini dibutuhkan informasi dan data yang mendukung kebenaran proses penelitian yang dilakukan. Maka penjelasan tahapan dan metode yang digunakan sesuai penelitian agar penulisan jurnal ini dapat tersusun rapi dan baik. Dimana alur penelitian dijelaskan dalam Gambar 3.



Gambar 3. Alur Penelitian

Identifikasi Masalah. Tahapan ini untuk menjelaskan permasalahan yang digunakan sebagai acuan dasar pembahasan pada tahap selanjutnya. Permasalahan yang menjadi dasar penelitian ini merupakan bagaimana penerapan algoritma Adam dalam mengoptimasi LSTM dalam memprediksi kualitas udara berdasarkan ISPU.

Studi literatur. Studi literatur yang memiliki peranan pada pengkajian ini dengan menggunakan informasi serta data yang terhimpun dari berbagai asal yang relevan. Studi literatur yang terkumpul dalam penelitian ini seperti kualitas udara, indeks standar pencemar udara, LSTM dan algoritma ADAM yang membantu dalam memecahkan dan memberi solusi pada permasalahan dalam penelitian.

Akuisisi Data. Pada proses ini merupakan pengumpulan data yang berhasil dikumpulkan dan dihimpun menjadi satu dataset yang diambil dari data terbuka Jakarta (data.jakarta.co.id). Dalam dataset yang berhasil dikumpulkan, terdapat beberapa kategori variabel yang ada pada dataset seperti CO2, O3, PM10, NO2, SO2, kategori, dan tempat pemantauan. Data yang terhimpun dalam open data Jakarta terhitung mulai bulan Januari tahun 2015 hingga bulan juli tahun 2021.

Praproses Data. Dimana data yang telah diambil melewati proses pemisahan data menjadi *data frame* terpisah pada setiap parameter ISPU (PM10,SO2,CO,O3,NO2) dan dilakukan split window dalam penentuan nilai fitur dan target yang digunakan dalam model LSTM dan Optimasi Adam, data akan mengalami praproses data yaitu dengan melakukan penyusunan dataset menjadi satu dan menghapus variabel pada kolom dan baris yang tidak digunakan serta normalisasi data. Pembagian data dilakukan pada data latih dan data uji yang mana masing-masing memiliki ratio 80% dan 20%. Dimana nilai ratio pada data training sebanyak 80% karena dalam proses *training* membutuhkan data yang banyak untuk memudahkan model LSTM dalam mempelajari pola *data time series* dalam proses pelatihan model sehingga pada proses pengujian data dapat memiliki akurasi yang baik.

Pembuatan Model. Pada perancangan model LSTM dan LSTMCell dari metode LSTM dengan model yang dibuat dalam memproses dalam memprediksi data parameter ISPU dengan melatih data uji dan data latih. Dimana data latih membutuhkan data yang banyak untuk menentukan pola nilai dari proses pelatihan metode LSTM. Pembuatan model dengan menyesuaikan arsitektur LSTM yang digunakan. Hiper Parameter yang digunakan dalam model LSTM adalah *learning rate*, *epoch*, *num layer*, dan *hidden size*. Sedangkan pada hiper parameter LSTMCell seperti *epoch*, *batch size*, dan *learning rate* digunakan sebagai pemodelan dari metode LSTM dalam pengolahan data.

Pengujian Model. Pengujian dilakukan pada data latih dan data uji. Pengujian dilakukan dengan menerapkan model LSTM yang belum dilakukan optimasi dengan menggunakan Adam dan pengujian terhadap model LSTM yang telah dioptimasi dengan Adam dalam meningkatkan hasil akurasi dalam model yang digunakan. Kemudian akan dilakukan perbandingan akurasi dari varian Adam dengan menggunakan model LSTMCell.

Evaluasi Model. Terdapat loss function yang digunakan sebagai kriteria dalam mengevaluasi kinerja prediksi kualitas udara termasuk *Root Mean Square Error* (RMSE), *R-square* (R²), dan *Mean Absolute Error* (MAE). Pada metode yang diterapkan dapat dikatakan bekerja dengan baik saat nilai *loss function* memiliki nilai yang kecil.

Perhitungan Matematis MAE yang dijelaskan dalam persamaan berikut :

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \dots \dots \dots (7)$$

Dimana \hat{y}_i merupakan nilai prediksi dan y_i merupakan nilai benar. Semakin kecil nilai dari MAE maka semakin baik kinerja model prediksinya.

Perhitungan Matematis RMSE yang dijelaskan dalam persamaan berikut :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \dots \dots \dots (8)$$

Dimana \hat{y}_i merupakan nilai prediksi dan y_i merupakan nilai benar. Semakin kecil nilai dari RMSE maka semakin baik kinerja model prediksinya.

Perhitungan Matematis R² yang dijelaskan dalam persamaan berikut :

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{(\sum_{i=1}^n (y_i - \bar{y})^2)} \dots \dots \dots (9)$$

Dimana \hat{y}_i merupakan nilai prediksi, y_i merupakan nilai benar, \bar{y} merupakan nilai rata-rata. Rentang nilai dari R² adalah (0,1).

Akurasi Model. Akurasi model untuk menentukan performa yang baik dari Adam menggunakan metode *Mean Absolute Percentage Error* (MAPE) dengan menemukan nilai kesalahan absolut dari setiap periode yang telah dibagi dengan nilai aktual dari periode tersebut dan dibuat rata-rata dari absolute percentage error.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \times 100 \dots\dots\dots (10)$$

Tabel 2. Akurasi MAPE

Nilai MAPE	Akurasi
≤ 10%	Sangat Akurat
11%-20%	Akurat
21%-50%	Cukup Akurat
>50%	Tidak Akurat

4 Hasil dan Pembahasan

Tabel 3 merupakan tabel yang akan digunakan dalam pemrosesan model dimana data tersebut nantinya akan mengalami praproses data terlebih dahulu dengan memisahkan setiap paramaternya menjadi bentuk dataframe baru untuk memudahkan model LSTM dalam mempelajari pola data yang dimasukan kedalam model sehingga menghasilkan output nilai prediksi pada setiap parameter dan perhitungan keakuratan model dari LSTM dengan penggunaan optimasi untuk meningkatkan hasil akurasi model.

Tabel 3. Dataset ISPU.

tanggal	pm10	so2	co	o3	no2	max	kategori
2015-01-01 00:00:00	49	10	39	27	10	49	BAIK
2015-01-02 00:00:00	28	9	26	48	10	48	BAIK
2015-01-03 00:00:00	27	10	28	44	10	44	BAIK
2015-01-04 00:00:00	22	9	28	42	12	42	BAIK
2015-01-05 00:00:00	25	9	25	32	7	32	BAIK
...
2021-07-27 00:00:00	82	56	13	41	35	140	TIDAK SEHAT
2021-07-28 00:00:00	82	53	18	40	45	145	TIDAK SEHAT
2021-07-29 00:00:00	78	52	18	53	39	140	TIDAK SEHAT
2021-07-30 00:00:00	90	54	15	81	35	154	TIDAK SEHAT
2021-07-31 00:00:00	63	50	18	65	29	100	SEDANG

Tahap praproses data dilakukan sebagai tahap awal dalam pengolahan data dan proses penelitian ini. Dalam praproses data setiap parameter dalam ISPU seperti PM10, SO2, CO, O3, dan NO2 akan dibuat menjadi *data frame* yang berbeda dimana berisi atribut tanggal dan masing-masing dari setiap parameter ISPU. Data yang telah mengalami proses *cleaning* dimana terdapat beberapa penghapusan fitur dari dataset udara DKI Jakarta seperti fitur max, critical, kategori, dan lokasi *spku*. Setiap *data frame* akan dilakukan normalisasi data yang dapat menimbulkan permasalahan dari model *machine learning* yang digunakan. *Feature scaling* digunakan agar nilai sama pada rentang nilai tertentu, rentang nilai yang digunakan antara -1 hingga 1 dengan menggunakan *feature scaling* MinMax Scaler. *Dataframe* yang telah mengalami normalisasi data akan terbagi menjadi dua yaitu data training dan data testing. Data *training* digunakan sebagai data untuk pelatihan algoritma *deep learning* sehingga mencapai model yang diinginkan.

Data *testing* digunakan sebagai gambaran mengenai kebenaran dan performa dari algoritma yang dilatih dalam model. Pembagian data tersebut dengan rasio pembagian 80% untuk data *training* dan 20% data *testing*. Diperoleh data *training* sebanyak 1907 data dari tahun 2015-01-01 hingga 2020-04-10 dan data *testing* sebanyak 477 data dari tahun 2020-04-11 hingga 2021-07-31. Pembagian data ini dilakukan pada data terurut yang dilakukan dari awal data hingga mencapai 80% untuk data *training* dan sisa dari data tersebut digunakan sebagai data *testing*.

Dimana data *training* lebih besar untuk memproses model lebih akurat karena membutuhkan data yang banyak dalam pemrosesan pembelajaran model sehingga model mampu menemukan pola dari data *time series* untuk memprediksi nilai pada setiap parameter ISPU sehingga dihasilkan nilai yang akurat. Pengujian dilakukan pada setiap parameter ISPU dengan menggunakan model LSTM dan optimasi Adam untuk memprediksi nilai pada setiap parameter dengan jangkauan waktu untuk 30 hari kedepan. Model LSTM menerima hiper parameter dengan learning rate sebesar 0,01, hidden size sebesar 120, num layer sebesar 3, dan epoch sebanyak 100.

Tabel 4. Hasil Akurasi MAPE pada model LSTM

Optimasi	MAPE					Rata-rata
	pm10	so2	co	o3	no2	
Tanpa Adam	18,97%	16,473%	42,193%	15,814%	59,44%	30,578%
Adam	4,37%	5,02%	18,50%	5,23%	37,28%	14,08%

Dijelaskan dalam Tabel 5. mengenai hasil evaluasi model LSTMCell pada parameter SO2 dengan menggunakan beberapa optimasi bahwa optimasi Adamax memiliki nilai terendah untuk RMSE, optimasi Adamax memiliki nilai terendah pada MAE, dan optimasi Adamax memiliki nilai tertinggi untuk nilai R².

Tabel 5. Hasil evaluasi performa model pada parameter ISPU

Optimasi	Evaluasi	pm10	so2	co	o3	no2
Adam	RMSE	11,0888	10,3582	6,0383	20,1528	22,1636
	MAE	8,5405	6,2146	4,3107	14,8510	14,1568
	R2	0,4438	0,8060	0,6453	0,4810	0,4543
AdamW	RMSE	11,1534	10,6113	5,8723	20,1159	24,5907
	MAE	8,5281	6,3028	4,1388	14,6816	15,8811
	R2	0,4373	0,7965	0,6646	0,4828	0,3282
Adamax	RMSE	11,0993	10,2149	6,0176	20,0501	22,4307
	MAE	8,5413	6,1391	4,2407	14,7613	14,3228
	R2	0,4427	0,8114	0,6478	0,4863	0,4411
NAdam	RMSE	11,0935	11,1765	5,9775	20,3490	21,8627
	MAE	8,5435	6,7069	4,2183	15,1619	13,8825
	R2	0,4433	0,7742	0,6525	0,4708	0,4691
RAdam	RMSE	11,0970	10,4413	5,9920	20,1780	22,4019
	MAE	8,5439	6,2532	4,2214	14,9055	14,2936
	R2	0,4430	0,8030	0,6508	0,4797	0,4425

Nilai MAPE yang kecil menunjukkan bahwa nilai yang dihasilkan antara prediksi pada model LSTMCell dan data aktual yang digunakan memiliki korelasi nilai yang dekat, hasil prediksi LSTMCell mampu menjelaskan probabilitas pada data aktual. Dengan penggunaan model LSTMCell dapat diketahui bahwa banyaknya hidden layer dan learning rate yang digunakan serta penambahan pada hiper parameter dalam proses pembangunan model memiliki peranan penting dalam proses pembelajaran model memiliki kontribusi besar terhadap proses prediksi LSTMCell. Distribusi data atau persebaran pada dataset sangat mempengaruhi nilai pada setiap parameter.

Tabel 6. Hasil evaluasi performa model pada parameter ISPU

Optimasi	PM10	SO2	CO	O3	NO2	Rata-Rata
Adam	15,68%	11,41%	24,74%	24,04%	30,61%	21,296%
AdamW	15,69%	11,37%	23,2%	23,43%	33,59%	21,456%
Adamax	15,69%	11,42%	23,65%	23,86%	30,63%	21,05%
NAdam	15,64%	11,73%	23,86%	24,87%	29,36%	21,092%

RAdam 15,68% 11,42% 23,7% 24,19% 30,46% 21,09%

Apabila dilihat dalam setiap parameter ISPU diketahui optimasi NAdam baik untuk digunakan prediksi parameter PM10 dengan nilai 15,64%, AdamW baik untuk digunakan dalam prediksi parameter SO2 dengan nilai 11,37%, pada parameter CO AdamW lebih unggul dari optimasi lainnya dengan nilai 23,2%, untuk parameter O3 memiliki nilai terendah pada 23,43% dimana AdamW lebih unggul, dan parameter NO2 memiliki nilai 29,36% pada optimasi NAdam. Dibandingkan dengan Adam maupun optimasi lainnya, nilai dari AdamW memiliki nilai yang baik dibandingkan dengan optimasi lainnya untuk prediksi pada setiap parameter dari ISPU.

Apabila nilai rata-rata MAPE pada setiap prediksi parameter bahwa nilai Adam lebih unggul dibandingkan nilai AdamW, dimana nilai untuk Adam 21,296% dan nilai AdamW yaitu 21,456%. Secara keseluruhan nilai rata-rata terkecil untuk tingkat akurasi MAPE dari penggunaan LSTM pada prediksi setiap parameter dengan menggunakan optimasi Adamax dengan nilai rata-rata MAPE 21,05%. Dibandingkan nilai NAdam, RAdam, dan Adamax, dimana Adam dan AdamW memiliki performa yang kurang dari NAdam, RAdam, dan Adamax yang merupakan pengembangan optimasi dari Adam sehingga dapat meningkatkan performa dari proses pembelajaran mesin dalam sebuah data time series dengan menggunakan model Long Short Term Memory.

5 Kesimpulan

Hasil kesimpulan dari pengujian yang telah dilakukan dalam proses memprediksi lima parameter dari ISPU (PM10, SO2, CO, O3, dan NO2) dengan menggunakan model *Long Short Term Memory* menghasilkan beberapa kesimpulan sebagai berikut:

1. Dengan pengujian pada parameter pada ISPU dengan menggunakan optimasi Adam. LSTM mampu memprediksi dengan baik pada 30 hari kedepan. Optimalnya nilai yang baik pada tujuh hari. Dengan penambahan data aktual akan berpengaruh dalam perubahan nilai prediksi untuk hari selanjutnya. Nilai akurasi MAPE dengan menggunakan model LSTM dan optimasi Adam diperoleh nilai pada parameter PM10 sebesar 4,37%, parameter SO2 sebesar 5,02%, parameter CO sebesar 18,50%, parameter O3 sebesar 5,23%, dan parameter NO2 sebesar 37,28%. Tanpa menggunakan Adam pada parameter PM10 sebesar 18,97%, parameter SO2 sebesar 16,473%, parameter CO sebesar 42,193%, parameter O3 sebesar 15,814%, dan parameter NO2 sebesar 59,44%. Dimana Adam melakukan optimasi pada nilai hidden state dari pemrosesan model LSTM setelah proses epoch awal dengan hasil proses epoch awal merupakan sebagai input awal pemrosesan.
2. Metode Long Short Term Memory dipakai dalam memprediksi lima parameter ISPU dengan model LSTMCell memiliki performa yang akurat dalam memprediksi pada setiap parameter dimana so2 memiliki nilai MAPE yang baik dengan nilai 11,37% pada optimasi AdamW, pm10 memiliki nilai MAPE 15,65% pada optimasi NAdam, co memiliki nilai MAPE 23,2% pada AdamW, o3 memiliki nilai MAPE 23,43% pada optimasi AdamW, dan no2 memiliki nilai MAPE 29,36% pada optimasi NAdam.
3. Jika dilakukan perbandingan pada setiap optimasi dimana nilai rata-rata MAPE dari kelima parameter ISPU dengan keseluruhan performa dari model untuk prediksi menghasilkan nilai 21,05% pada optimasi Adamax lebih unggul dari optimasi Adam pada 21,29%, dilanjutkan dengan nilai RAdam sebesar 21,09%, nilai NAdam sebesar 21,09%, dan tertinggi pada AdamW sebesar 21,45%. Apabila optimasi digunakan dalam menilai performa model dari setiap parameter, optimasi AdamW lebih unggul pada tiga parameter yaitu so2, co, dan o3. Sedangkan pada pm10 dan no2 lebih unggul menggunakan NAdam.

6 Daftar Pustaka

- [1] Aldi, M. P., Jondri, & Aditsania, A. (2018). Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin. *e-Proceeding of Engineering*.
- [2] Aprianto, Y., Nurhasanah, & Sanubary, I. (2018). Prediksi Kadar Particulate Matter (PM10) untuk Pemantauan Kualitas Udara Menggunakan Jaringan Syaraf Tiruan Studi Kasus Kota Pontianak. *POSITRON*.
- [3] Ba, J. L., & Kingma, D. P. (2015). ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION.

- [4] Chandra, R., Gupta, R., & Tiwaria, A. (2021). Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown. *arXiv*.
- [5] Chandriah, K. K., & Naraganahalli, R. V. (2021). RNN / LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting. *Springer Nature*.
- [6] Chang, Z., Zhang, Y., & Chen, W. (2018). Effective Adam-optimized LSTM Neural Network for Electricity Price Forecasting. *IEEE*, 245.
- [7] Elampartihi, P. N., Janarthanan, R., Partheeban, P., & Somasundaram, K. (2021). A deep learning approach for prediction of air quality index in a metropolitan city. *ELSEVIER*.
- [8] Graves, A. (2012). *Supervised Sequence Labell with RNN*. Springer.
- [9] Gul, S., & Khan, G. M. (2020). Forecasting Hazard Level of Air Pollutants Using LSTM's. *Springer Nature Switzerland*, 143.
- [10] HE, H., & LUO, F. (2020). Study of LSTM Air Quality Index Prediction Based on Forecasting Timeliness. *IOP Conference Series: Earth and Environmental Science*.
- [11] He, J., Jiang, F., & Tian, T. (2019). A clustering-based ensemble approach with improved pigeon-inspired optimization and extreme learning machine for air quality prediction. *Applied Soft Computing Journ*.
- [12] Jiao1, Y., Wang, Z., & Zhang, Y. (2019). Prediction of Air Quality Index Based on LSTM . *Joint International Information Technology and Artificial Intelligence Conference*, 8, 17.
- [13] Kementerian Lingkungan Hidup. (1997). *Standar Keputusan ISPU*. MENTERI LINGKUNGAN HIDUP DAN KEHUTANAN REPUBLIK INDONESIA.
- [14] Khumadi, A., Raafi'udin, R., & Solihin, I. (2019). Pengujian Algoritma Long Short Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung. *Jurnal Telematika*.
- [15] Kumari, S., Chaudhry, I., Sharma, S., & Sethi, P. (2019). AQI: PREDICTION AND OPTIMIZATION . *TECHNIQUES*, 6(6), 675.
- [16] Li, H., Wang, J., Li, R., & Lu, H. (2018). Novel analysis–forecast system based on multi-objective optimization for air quality index. *ournal of Cleaner Production*.
- [17] Liang, Y.-C., Maimury, Y., Chen, A. H.-L., & Juarez, J. R. (2020). Machine Learning-Based Prediction of Air Quality. *10*.
- [18] Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-Based Model to Forecast Stock Prices. *Hindawi*, 10.
- [19] Meliana1, C., Wasono, R., & Haris, M. A. (2020). Perbandingan Metode Long Short Term Memory (LSTM) DAN Genetic Algorithm-Long Short Term Memory (GA-LSTM) Pada Peramalan Polutan Udara.
- [20] Mukherjee, P., & Roy, S. (2020). AIR QUALITY INDEX FORECASTING USING HYBRID NEURAL NETWORK MODEL WITH LSTM ON AQI SEQUENCES. *Proceedings on Engineering Sciences*.
- [21] Muslim, B., & Prabowo, K. (2018). *Penyehatan Udara*. Pusat Pendidikan Sumber Daya Kesehatan.
- [22] Oktaviani, A., & Hustinawati. (2019). PREDIKSI RATA-RATA ZAT BERBAHAYA DI DKI JAKARTA BERDASARKAN INDEKS STANDAR PENCEMAR UDARA MENGGUNAKAN METODE LONG SHORT-TERM MEMORY.
- [23] R, N., Bhumika, S., R, S., & V, R. (2020). Air Quality Index Prediction using LSTM. *International Research Journal of Engineering and Technology (IRJET)*, 7, 4848.
- [24] Saurabh, N. (2020). LSTM -RNN Model to Predict Future Stock Prices using an Efficient Optimizer. *International Research Journal of Engineering and Technology (IRJET)*, 7(11), 672.
- [25] Wright, S. J. (2016). optimization.
- [26] Xayasouk, T., Lee, H., & Lee, G. (2020). Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models. *Sustainability*, 12, 1-18.
- [27] Xu, Y., Zhang, D., & Zhao, Q. (2019). Prediction of Air Quality Index Based on LSTM Model: A Case Study on Delhi and Houston. *Journal of Computer Research and Development*.