

# PERAMALAN KUALITAS UDARA DI KOTA JAKARTA PUSAT DENGAN METODE LONG SHORT-TERM MEMORY DAN SUPPORT-VECTOR REGRESSION

Mohammad Ernico Suryo Wicaksono<sup>a1</sup>, Gusti Made Arya Sasmita<sup>a2</sup>, I Putu Agus Eka Pratama<sup>b3</sup>

<sup>a</sup>Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana, Bali, Indonesia  
e-mail: <sup>1</sup>ernico@student.unud.ac.id, <sup>2</sup>aryasasmita@unud.ac.id, <sup>3</sup>eka.pratama@unud.ac.id

## Abstrak

Kualitas udara merupakan masalah serius di kota-kota besar di seluruh dunia, karena memiliki dampak yang signifikan terhadap kesehatan masyarakat dan lingkungan. Penelitian ini membahas mengenai peramalan kualitas udara di Kota Jakarta Pusat menggunakan Metode Long Short-Term Memory (LSTM) dan Support-Vector Regression (SVR). Tujuan dari penelitian ini untuk membantu para pembuat kebijakan dalam membuat keputusan tentang pengelolaan dan penanggulangan kualitas udara serta dalam mengambil tindakan yang sesuai ke depannya untuk melindungi kesehatan masyarakat dan lingkungan. Parameter yang diramalkan dengan Metode LSTM dan SVR adalah nilai konsentrasi polutan Particulate Matter 10 (PM<sub>10</sub>), Sulfur Dioksida (SO<sub>2</sub>), Karbon Monoksida (CO), Ozon Permukaan (O<sub>3</sub>), dan Nitrogen Dioksida (NO<sub>2</sub>) dengan evaluasi menggunakan metrik Mean Absolute Percentage Error (MAPE) dan Root-Mean-Square Error (RMSE). Hasil yang didapat dari Metode LSTM adalah rata-rata nilai MAPE sebesar 12,15% dan rata-rata nilai RMSE sebesar 0,0941, sedangkan Metode SVR memiliki hasil rata-rata nilai MAPE sebesar 16,19% dan rata-rata nilai RMSE sebesar 0,1666.

**Kata kunci:** Kualitas Udara, Polusi Udara, Data Mining, Long Short-Term Memory, Support-Vector Regression

## Abstract

Air quality is a critical issue in many major cities around the world, as it has significant implications for public health and the environment. This research proposed air quality forecasting at Central Jakarta using Long Short-Term Memory (LSTM) and Support-Vector Regression (SVR) methods. The purpose of this research is to help policymakers to make decisions about air quality management and take appropriate actions to protect public health and the environment. The forecasted parameters are Particulate Matter 10 (PM<sub>10</sub>), Sulfur Dioxide (SO<sub>2</sub>), Carbon Monoxide (CO), Ground-Level Ozon (O<sub>3</sub>), and Nitrogen Dioxide (NO<sub>2</sub>) using Mean Absolute Percentage Error (MAPE) and Root-Mean-Square Error (RMSE) as its evaluation metrics. The results showed that LSTM method has an average of 12.15% MAPE value and 0.0941 RMSE value, while SVR method has an average of 16.19% MAPE value and 0.1666 RMSE value.

**Keywords :** Air Quality, Air Pollution, Data Mining, Long Short-Term Memory, Support-Vector Regression

## 1. Pendahuluan

Kota Administrasi Jakarta Pusat atau lebih dikenal dengan Kota Jakarta Pusat adalah kota administrasi yang terletak di Provinsi Daerah Khusus Ibukota (DKI) Jakarta dengan luas 48,13 km<sup>2</sup> dengan populasi penduduk berdasarkan data kependudukan tahun 2021 dari Badan Pusat Statistik (BPS) mencapai 1.136.824 jiwa [1]. Polusi udara di Provinsi DKI Jakarta pada umumnya dan Kota Jakarta Pusat khususnya terus meningkat, salah satunya dikarenakan

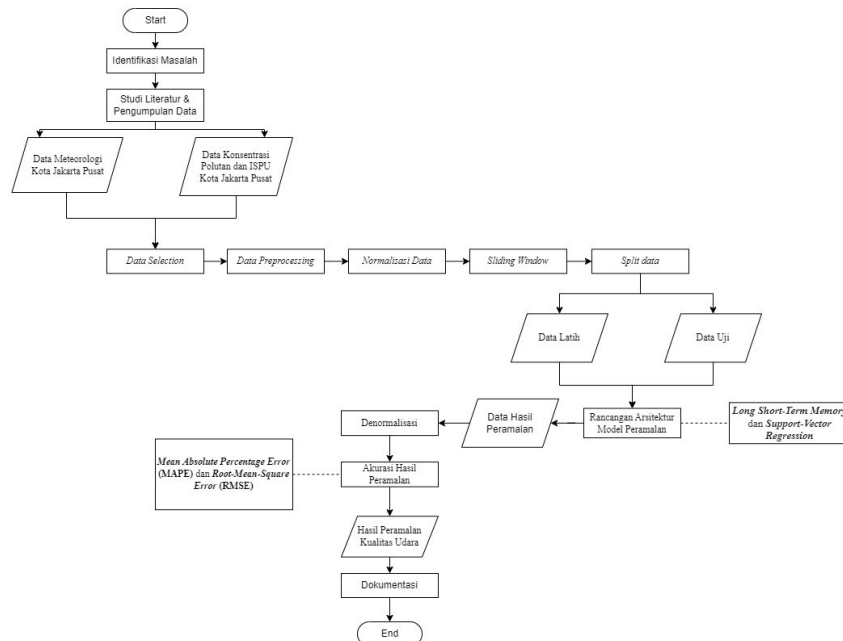
---

penambahan kendaraan bermotor yang mayoritas masih berbahan bakar fosil dengan penambahan rata-rata setiap tahunnya sebanyak 12% dan didominasi oleh sepeda motor pribadi dengan rata-rata penambahan sebanyak 73% dibandingkan jenis kendaraan bermotor lainnya [2]. Polusi udara menyebabkan kualitas udara memburuk sehingga memengaruhi kesehatan masyarakat, terutama anak-anak. Terbukti dari 60 orang anak yang berkegiatan di luar ruangan dalam kurun waktu 1-6 jam, 58 orang anak diantaranya mengalami batuk, 54 orang anak mengalami sakit tenggorokan, 43 orang anak mengalami sesak, 40 orang anak mengalami asma, 37 orang anak mengalami pilek, dan 36 orang anak mengalami nyeri [3].

Permasalahan udara di Kota Jakarta Pusat dapat diantisipasi salah satunya dengan peramalan kualitas udara di Kota Jakarta Pusat, dimana salah satu tujuannya adalah untuk membantu dalam pengambilan keputusan dan penetapan kebijakan oleh pemangku kebijakan dalam mengatasi permasalahan yang timbul akibat polusi udara untuk memperbaiki kualitas udara di Kota Jakarta Pusat sehingga kualitas udara dimasa mendatang menjadi lebih baik. Penelitian ini menggunakan Metode *Long Short-Term Memory* (LSTM) dan *Support-Vector Regression* (SVR) untuk melakukan peramalan dengan tujuan untuk mengetahui metode mana yang lebih optimal dalam peramalan kualitas udara di Kota Jakarta Pusat diantara kedua metode tersebut. Akurasi dari hasil peramalan pada masing-masing metode dihitung dengan menggunakan *Mean Absolute Percentage Errors* (MAPE) dan *Root-Mean-Square Error* (RMSE) untuk menentukan metode peramalan mana yang lebih baik diantara kedua metode tersebut dalam meramalkan kualitas udara di Kota Jakarta Pusat.

## 2. Metodologi Penelitian

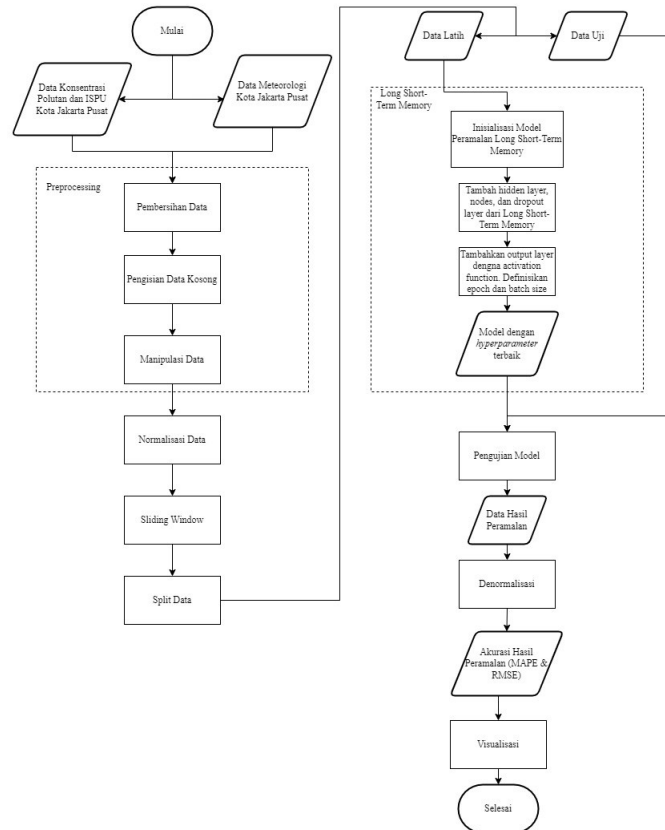
Penelitian ini menggunakan data rata-rata *Particulate Matter* 10 (PM<sub>10</sub>), Sulfur Dioksida (SO<sub>2</sub>), Karbon Monoksida (CO), Ozon Permukaan (O<sub>3</sub>), dan Nitrogen Dioksida (NO<sub>2</sub>) di Kota Jakarta Pusat yang diperoleh dari Dinas Lingkungan Hidup Provinsi DKI Jakarta melalui Portal Data Terpadu Pemerintah Provinsi DKI Jakarta dan data meteorologi yang memengaruhi kelima parameter tersebut seperti data rata-rata temperatur, kelembapan, curah hujan, dan kecepatan angin di Kota Jakarta Pusat yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) melalui Aplikasi Data Online - Pusat Database - BMKG. Rentang waktu dari data yang digunakan dari Januari 2010 hingga Desember 2019. Gambaran umum penelitian ini digambarkan sebagai berikut.



Gambar 1. Gambaran Umum Penelitian

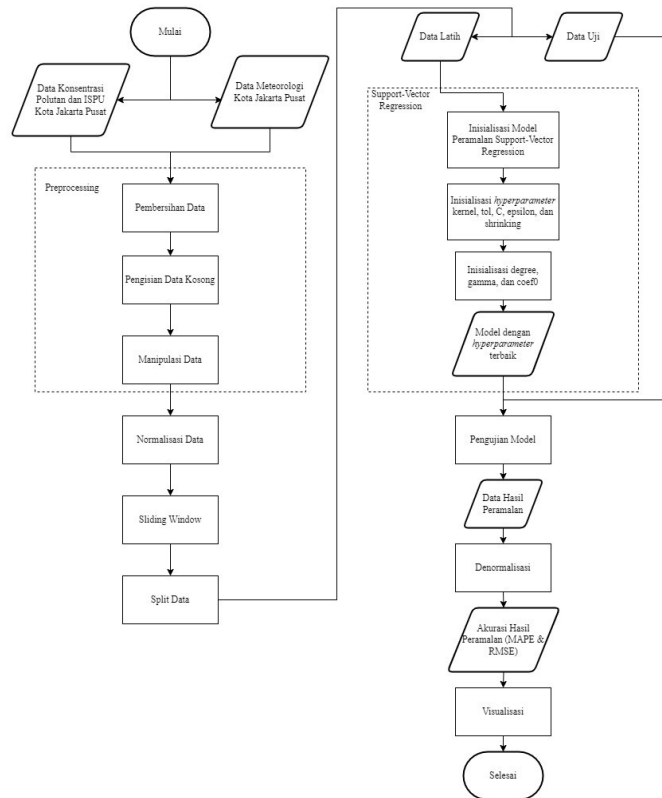
Gambar 1. menjelaskan gambaran umum penelitian ini. Tahap pertama adalah identifikasi masalah, kemudian dilanjutkan dengan tahap studi literatur dan pengumpulan data,

kemudian dilanjutkan dengan memroses data dengan *data selection*, *data preprocessing*, normalisasi data, *sliding window*, *split data* menjadi data latih dan data uji, merancang arsitektur model peramalan dengan Metode LSTM dan SVR, denormalisasi hasil peramalan, menghitung akurasi hasil peramalan dengan MAPE dan RMSE, dan terakhir adalah mendokumentasikan hasil. Metode peramalan yang digunakan memiliki alur peramalan berbeda-beda. Alur peramalan Metode LSTM digambarkan sebagai berikut.



Gambar 2. Alur Peramalan *Long Short-Term Memory* (LSTM)

Gambar 2. merupakan alur peramalan LSTM yang dimulai dengan memroses data terlebih dahulu agar data dapat digunakan untuk proses peramalan dengan LSTM. Arsitektur model diinisialisasi terlebih dahulu, kemudian mendefinisikan *hyperparameter*-nya seperti *hidden layer*, *nodes* atau *neuron*, *dropout*, *output layer*, *activation function*, *epoch*, dan *batch size*. Arsitektur model yang sudah dibuat beserta nilai *hyperparameters*-nya kemudian di latih dengan data latih, kemudian diuji dengan data uji sehingga didapat hasil yang diproses dengan denormalisasi agar didapat nilai aktualnya. Alur peramalan Metode SVR digambarkan sebagai berikut.



Gambar 3. Alur Peramalan *Support-Vector Regression* (SVR)

Gambar 3. merupakan alur peramalan SVR yang dimulai dengan memroses data terlebih dahulu agar data dapat digunakan untuk proses peramalan dengan SVR. Arsitektur model diinisialisasi terlebih dahulu, kemudian mendefinisikan *hyperparameter*-nya seperti *kernel*, *tol*, *C*, *epsilon*, *shrinking*, *degree* (*kernel Polynomial*), *gamma* (*kernel RBF* atau *Polynomial*), dan *coef0* (*kernel Polynomial*). Arsitektur kedua model yang sudah dibuat beserta nilai *hyperparameters*-nya kemudian di latih dengan data latih, kemudian diuji dengan data uji sehingga didapat hasil yang diproses dengan denormalisasi agar didapat nilai aktualnya. Akurasi hasil dari kedua metode dihitung dengan MAPE dan RMSE dan langkah terakhir adalah data divisualisasikan.

### 2.1 Mean Absolute Percentage Error (MAPE)

*Mean Absolute Percentage Errors* (MAPE) merupakan ukuran kesalahan relatif. MAPE menyatakan persentase kesalahan hasil peramalan terhadap permintaan aktual selama periode tertentu yang memberikan informasi persentase kesalahan terlalu tinggi atau terlalu rendah. Persamaan MAPE adalah sebagai berikut [4] :

$$MAPE = \frac{\sum \frac{|X_i - F_i|}{X_i}}{n} 100\% \quad (1)$$

Parameter  $X_i$  menunjukkan nilai aktual, parameter  $F_i$  merupakan nilai peramalan, dan parameter  $N$  adalah jumlah data. Tabel 1. menunjukkan tingkat akurasi hasil peramalan berdasarkan nilai MAPE.

Tabel 1. Akurasi Hasil Berdasarkan MAPE

MAPE	Hasil Peramalan
<10%	Sangat Baik
10-20%	Baik

20-50%	Layak/Cukup
>50%	Buruk

Tabel 1. menunjukkan tingkat akurasi hasil peramalan berdasarkan nilai MAPE. Sebuah model peramalan jika berhasil menghasilkan nilai MAPE <10%, maka model tersebut dapat dikategorikan sangat akurat.

### 2.2 Root-Mean-Square Error (RMSE)

Root-Mean-Square Error (RMSE) merupakan akar kuadrat estimasi perbedaan nilai peramalan dengan nilai aktual. Semakin tinggi nilai RMSE semakin akurat nilai yang dihasilkan model peramalannya. Persamaan RMSE adalah sebagai berikut [5] :

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(P_m - P_r)^2}{N}} \quad (2)$$

Parameter  $P_m$  berarti nilai aktual, parameter  $P_r$  menunjukkan nilai peramalan, dan parameter  $N$  adalah jumlah data.

## 3. Kajian Pustaka

Kajian pustaka merupakan tahap mengumpulkan dan mempelajari teori terkait dengan permasalahan yang diangkat dalam penelitian ini. Kajian pustaka yang dijabarkan diantaranya polusi udara, Indeks Standar Pencemar Udara (ISPU), normalisasi data, *sliding window*, peramalan, *Long Short-Term Memory* (LSTM), dan *Support-Vector Regression* (SVR).

### 3.1 Polusi Udara

Polusi udara atau pencemaran udara adalah tercampurnya substansi, baik fisik, kimia, ataupun biologi, di udara bersih dalam jumlah yang membahayakan kesehatan makhluk hidup, merusak lingkungan, hingga bisa bersifat korosif terhadap properti [6].

### 3.2 Indeks Standar Pencemar Udara (ISPU)

Peraturan Menteri Lingkungan Hidup dan Kehutanan (Permen LHK) Nomor 14 Tahun 2020 tentang Indeks Standar Pencemar Udara (ISPU) menyebutkan bahwa ISPU merupakan gambaran kualitas udara ambien di suatu lokasi dalam bentuk angka tanpa satuan yang didasarkan kepada kesehatan manusia dan makhluk hidup lainnya. Pasal 2 Ayat 2 pada Permen LHK Nomor 14 Tahun 2020 menyatakan bahwa polutan yang terdiri dari *Particulate Matter 2,5* ( $PM_{2.5}$ ), *Particulate Matter 10* ( $PM_{10}$ ), Sulfur Dioksida ( $SO_2$ ), Karbon Monoksida (CO), Ozon Permukaan ( $O_3$ ), Nitrogen Dioksida ( $NO_2$ ), dan Hidrokarbon (HC) diukur kemudian dihitung agar menjadi nilai ISPU.

#### 3.2.1 *Particulate Matter 10* ( $PM_{10}$ )

*Particulate Matter 10* ( $PM_{10}$ ) merupakan polutan udara berbentuk partikel yang terdiri partikel padat ataupun cair berukuran lebih kecil atau sama dengan 10 mikrometer dan lebih besar dari 2,5 mikrometer. Sumber polutan  $PM_{10}$  biasanya berasal dari konstruksi bangunan, tempat pembuangan sampah, pertanian, sampah industri, dan angin yang menghembuskan debu dari lapangan/ruang terbuka [7].

#### 3.2.2 Sulfur Dioksida ( $SO_2$ )

Sulfur Dioksida ( $SO_2$ ) adalah gas berbahaya yang biasanya dihasilkan dari hasil pembakaran menggunakan bahan bakar fosil yang mengandung sulfur, kendaraan diesel, dari aktivitas industri yang menggunakan bahan bakar minyak mentah dan batu bara, atau dari gunung berapi [8].

#### 3.2.3 Karbon Monoksida (CO)

Karbon Monoksida (CO) merupakan polutan udara berjenis gas yang tidak berwarna, tidak berbau, tidak berasa, dan tidak mengiritasi. CO biasanya bersumber dari hasil pembakaran tidak sempurna pada mesin berbahan bakar fosil, minyak tanah, bahan bakar padat, dan gas. Efek kesehatan yang dapat ditimbulkan bila seseorang mengalami keracunan

CO diantaranya sakit kepala, pusing, kelelahan, mual, muntah-muntah, hingga hilang kesadaran [9].

### 3.2.4 Ozon Permukaan (O<sub>3</sub>)

Ozon Permukaan (O<sub>3</sub>) adalah polutan udara sekunder yang terbentuk dari hasil reaksi kimiawi antara oksida nitrogen dan senyawa organik yang mudah menguap yang berasal baik dari sumber alami maupun dari aktivitas manusia. Kadar O<sub>3</sub> yang tinggi dapat membahayakan hutan dan tumbuh-tumbuhan serta mengurangi potensi hasil panen pertanian. Paparan terhadap konsentrasi O<sub>3</sub> yang tinggi memiliki dampak buruk bagi kesehatan, diantaranya kerusakan paru-paru, gangguan pernapasan, penyakit paru obstruktif kronik, peningkatan morbiditas, dan mortalitas [9].

### 3.2.5 Nitrogen Dioksida (NO<sub>2</sub>)

Nitrogen Dioksida (NO<sub>2</sub>) merupakan polutan udara yang terbentuk dari pembakaran pada temperatur tinggi dari bahan bakar fosil seperti batu bara, minyak mentah, bensin, ataupun diesel. Efek buruk terhadap kesehatan dari NO<sub>2</sub> adalah berbagai macam penyakit pernafasan, karena NO<sub>2</sub> yang terhirup masuk dan menembus hingga ke bagian dalam paru-paru, serta dapat memengaruhi kesehatan mata, tenggorokan dan menyebabkan iritasi pada hidung. Efek buruk terhadap lingkungan dari NO<sub>2</sub> adalah merusak tumbuh-tumbuhan [9].

### 3.3 Normalisasi Data

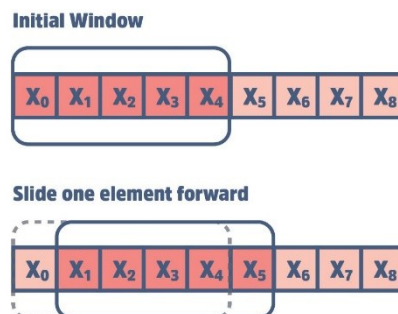
Normalisasi data adalah proses mengubah nilai aktual ke dalam nilai yang lebih mudah dipahami agar memudahkan proses pengolahan data, mengurangi redundansi data, membuat perhitungan numerik menjadi lebih tepat, dan meningkatkan akurasi hasil peramalan [10]. Persamaan yang digunakan pada normalisasi *min-max* adalah sebagai berikut [11] :

$$X'_i = \left( \frac{X_i - X_{min}}{X_{max} - X_{min}} \right) (X'_{max} - X'_{min}) + X'_{min} \quad (1)$$

Parameter X<sub>i</sub> merupakan nilai aktual, parameter X<sub>min</sub> menunjukkan nilai aktual terendah, parameter X<sub>max</sub> adalah nilai aktual tertinggi, parameter X'<sub>min</sub> menggambarkan nilai target terendah, dan X'<sub>max</sub> berarti nilai target tertinggi.

### 3.4 Sliding Window

*Sliding Window* adalah metode yang merubah data *time series* menjadi *supervised* dengan cara memperkirakan nilai sementara data *time series* menggunakan *time step* (langkah waktu) sebagai input X untuk memprediksi data pada langkah waktu selanjutnya. Parameter langkah waktu ini disebut dengan *window size*. Langkah waktu selanjutnya dijadikan sebagai parameter Y atau sebagai keluaran dari input X [12]. *Sliding Window* dapat digambarkan sebagai berikut.



Gambar 4. Ilustrasi *Sliding Window*

Gambar 4. merupakan ilustrasi dari *sliding window*. Gambar tersebut menjelaskan jika data yang dimiliki misalnya adalah data bulanan, maka *window size* sebesar 5 mengartikan bahwa data 5 bulan terakhir (input X) digunakan untuk memprediksi data pada langkah waktu

selanjutnya (keluaran Y) dan berlanjut hingga data *time series* yang digunakan pada periode waktu tertentu mencapai titik akhir.

### 3.5 Peramalan

Peramalan merupakan kegiatan memprediksi nilai, kejadian, peristiwa, waktu, lokasi, kuantitas, atau kualitas di masa mendatang yang didasarkan pada pertimbangan subjektif terhadap data masa lalu sehingga informasi yang dihasilkan dari peramalan dapat menjadi referensi penting bagi masyarakat [13]. Metode peramalan kuantitatif adalah metode peramalan dengan memanfaatkan informasi di masa lalu dan metode kualitatif adalah metode peramalan dengan memanfaatkan pendapat dari ahli atau pakar untuk pengambilan keputusan dari hasil peramalan yang dilakukan [14].

### 3.6 Long Short-Term Memory (LSTM)

*Long Short-Term Memory* (LSTM) adalah metode yang pertama kali diperkenalkan oleh Hochreiter & Schmidhuber pada tahun 1997 dan merupakan hasil pengembangan dari arsitektur *Recurrent Neural Network* (RNN) [15]. LSTM pada prosesnya melibatkan empat komponen utama, diantaranya *forget gate*, *input gate*, *cell state*, dan *output gate*. *Forget gate* berfungsi untuk membedakan data yang penting dan tidak penting pada yang dijalankan dengan fungsi *sigmoid*. *Input gate* bertugas mengendalikan seberapa banyak informasi terbaru yang didapat yang akan lewat. *Cell state* memiliki fungsi untuk menyimpan informasi sebelumnya. *Output gate* berfungsi untuk menentukan informasi mana yang hendak dikeluarkan dari *cell state* [5].

### 3.7 Support-Vector Regression (SVR)

*Support-Vector Regression* (SVR) adalah metode *machine learning* (pembelajaran mesin) yang merupakan bagian dari *Support-Vector Machine* (SVM) dan juga dikenal sebagai *SVM-Regression*. SVR dikhususkan untuk memecahkan kasus regresi dan memberikan *output* dalam bentuk data yang kontinu dengan bilangan riil sehingga dapat digunakan untuk melakukan peramalan. Kelebihan yang dimiliki oleh Metode SVR adalah metode ini merupakan salah satu metode yang digunakan untuk mengatasi *overfitting* dan mampu menangkap tren data yang non-linier sehingga hasil peramalan lebih akurat [16].

## 4. Hasil dan Pembahasan

Penelitian ini berfokus pada komparasi Metode LSTM dan SVR dalam meramalkan rata-rata PM<sub>10</sub>, SO<sub>2</sub>, CO, O<sub>3</sub>, dan NO<sub>2</sub>. Contoh data kualitas udara yang sudah dinormalisasi digambarkan sebagai berikut.

Tabel 2. Contoh Data Kualitas Udara

Bulan	PM <sub>10</sub> (μ g/m <sup>3</sup> )	PM <sub>10</sub> ISPU	SO <sub>2</sub> (μ g/m <sup>3</sup> )	SO <sub>2</sub> ISPU	CO (μ g/m <sup>3</sup> )	CO ISPU	O <sub>3</sub> (μ g/m <sup>3</sup> )	O <sub>3</sub> ISPU	NO <sub>2</sub> (μ g/m <sup>3</sup> )	NO <sub>2</sub> ISPU
Jan 2010	0.180 481	0.270 833	0.079 082	0.055 556	0.639 05	0.625	0.115 162	0.111 111	0.548 851	0.523 81
Feb 2010	0.371 658	0.5	0.112 245	0.111 111	0.876 092	0.875	0.381 944	0.375	0.747 126	0.761 905
Mar 2010	0.413 102	0.541 667	0.163 265	0.166 667	0.913 698	0.875	0.405 093	0.388 889	0.761 494	0.761 905
Apr 2010	0.415 775	0.541 667	0.275 51	0.277 778	0.820 911	0.791 667	0.504 63	0.486 111	0.701 149	0.714 286
Mei 2010	0.491 979	0.604 167	0.375	0.361 111	0.919 086	0.916 667	0.539 352	0.527 778	0.816 092	0.809 524

Tabel 2. merupakan contoh data kualitas udara yang sudah dinormalisasi. Rentang waktu dari data yang digunakan adalah Januari 2010 hingga Desember 2019. Data meteorologi yang sudah dinormalisasi digambarkan sebagai berikut.

Tabel 3. Contoh Data Meteorologi

Bulan	temperatur_rata-rata(°C)	kelembapan_rata-rata(%)	curah_hujan (mm)	kecepatan_angin_rata-rata(m/s)
Jan 2010	0.242424	0.751351	0.436364	0.242424
Feb 2010	0.484848	0.756757	0.290909	0.484848
Mar 2010	0.606061	0.659459	0.284848	0.606061
Apr 2010	0.909091	0.443243	0.066667	0.909091
Mei 2010	0.818182	0.562162	0.084848	0.818182

Tabel 3. merupakan contoh data meteorologi yang sudah dinormalisasi. Rentang waktu dari data yang digunakan adalah Januari 2010 hingga Desember 2019.

#### 4.1. Hasil Peramalan dengan Long Short-Term Memory (LSTM)

Hasil terbaik yang didapat dari Metode LSTM pada penelitian ini menggunakan *window size* sebesar 12 dan skema peramalan 9 tahun data latih dan 1 tahun data uji. Hasil terbaik peramalan dari masing-masing parameter dapat dilihat pada Tabel 4.

Tabel 4. Hasil Terbaik LSTM

Parameter	Hidden Layers	Neuron	Activation Function	Drop out	Learning Rate	Batch Size	Epochs	MAPE	RMSE
PM <sub>10</sub>	1	70	Linear	0,1	0,004	41	230	9,379 %	0,0994
SO <sub>2</sub>	2	10	Linear	False	0,01	6	500	18,356 %	0,0954
		100							
CO	2	8	Linear	False	0,001	60	500	7,376 %	0,0889
		128							
O <sub>3</sub>	1	100	Linear	False	0,0079665	1	500	11,362 %	0,1110
NO <sub>2</sub>	3	512	Linear	0,1	0,01	1	500	14,3%	0,0756
		8							
		60							

Tabel 4. merupakan hasil terbaik dari Metode LSTM dalam peramalan di setiap parameter. Hasil-hasil tersebut jika dirata-ratakan, maka rata-rata MAPE yang didapat adalah 12,15% dan rata-rata RMSE yang didapat adalah 0,0941. Hasil akhir peramalan dengan Metode LSTM setelah proses denormalisasi dapat dilihat pada Tabel 5.

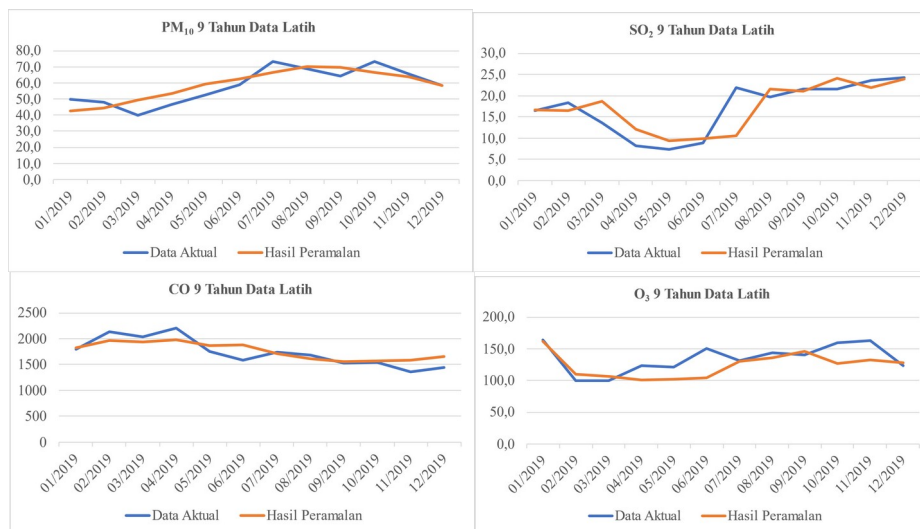
Tabel 5. Hasil Akhir Peramalan LSTM

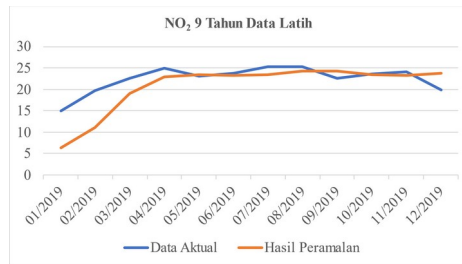
Bulan	Hasil Peramalan									
	PM <sub>10</sub> Data Aktual	PM <sub>10</sub> Hasil Peramalan	SO <sub>2</sub> Data Aktual	SO <sub>2</sub> Hasil Peramalan	CO Data Aktual	CO Hasil Peramalan	O <sub>3</sub> Data Aktual	O <sub>3</sub> Hasil Peramalan	NO <sub>2</sub> Data Aktual	NO <sub>2</sub> Hasil Peramalan
Januari 2019	49,8	42,34393018	16,5	16,71515086	1800	1822,366309	163,5	162,0937737	15	6,251018715
Februari 2019	47,9	44,28181242	18,4	16,46856928	2133,3	1970,795963	99,2	109,3729582	19,6	10,98579054
Maret 2019	39,8	49,31845553	13,6	18,6151628	2037,3	1933,080953	99,8	106,3867661	22,5	19,09040022



April 2019	46,5	53,486 66831	8,2	12,114 98482	2200	1982,9 15352	123, 3	100,4 18228 7	24,9	22,855 47652
Mei 2019	52,5	59,117 15835	7,4	9,4348 94276	1751 ,4	1864,0 19312	120, 6	101,5 82556 3	23,1	23,327 83942
Juni 2019	58,9	62,484 93629	8,8	9,9015 38718	1582 ,9	1877,5 65115	149, 9	103,8 84077 1	23,8	23,211 11977
Juli 2019	73,2	66,517 58835	21,8	10,533 07478	1741 ,5	1713,5 63426	131, 7	130,0 80423 7	25,3	23,440 031
Agus tus 2019	68,8	70,205 35431	19,7	21,491 23564	1680	1617,4 71147	143, 1	135,1 77674 9	25,3	24,259 34401
Sep- tem- ber 2019	64,3	69,722 02065	21,6	20,971 99011	1523 ,2	1554,5 74099	140, 4	145,9 44777 7	22,5	24,317 22987
Okto- ber 2019	73,4	66,307 15504	21,6	24,081 73418	1540	1563,5 27703	159, 5	127,0 56642 9	23,6	23,409 98354
No- vem- ber 2019	65,6	63,589 89733	23,5	21,834 7178	1363 ,2	1583,3 98239	162, 7	131,7 93168 3	24	23,283 37598
De- sem- ber 2019	58,3	58,257 36177	24,3	23,917 29584	1440	1649,5 41868	123, 3	127,6 62480 5	19,9	23,712 78548

Tabel 5. merupakan tabel perbandingan antara data aktual dari parameter-parameter yang diramalkan dengan hasil peramalan menggunakan Metode LSTM. Peramalan yang dilakukan adalah peramalan selama 12 bulan ke depan pada tahun 2019. Visualisasi hasil dapat dilihat pada Gambar 5.





Gambar 5. Visualisasi Hasil Peramalan LSTM

Gambar 5. merupakan visualisasi hasil peramalan parameter dengan Metode LSTM. Grafik biru merupakan data aktual, sedangkan grafik oranye merupakan data hasil peramalan.

**4.2. Peramalan dengan Support-Vector Regression (SVR)**

Hasil terbaik yang didapat dari Metode SVR pada penelitian ini menggunakan *window size* sebesar 3 dan skema peramalan 9 tahun data latih dan 1 tahun data uji. Hasil terbaik peramalan dari masing-masing parameter dapat dilihat pada Tabel 6.

Tabel 6. Hasil Terbaik SVR

kernel	gamma	tol	C	epsilon	shrinking	MAPE	RMSE
Linear	-	0,008	0,1	0,1006	False	10,348%	0,2017
Linear	-	0,0019	100	0,001	True	24,067%	0,158301
RBF	auto	0,005	10	0,1862	False	10,418%	0,15188
Linear	-	0,001	1	0,0906	False	17,308%	0,16021
Linear	-	0,006	1	0,0271	True	18,805%	0,1610

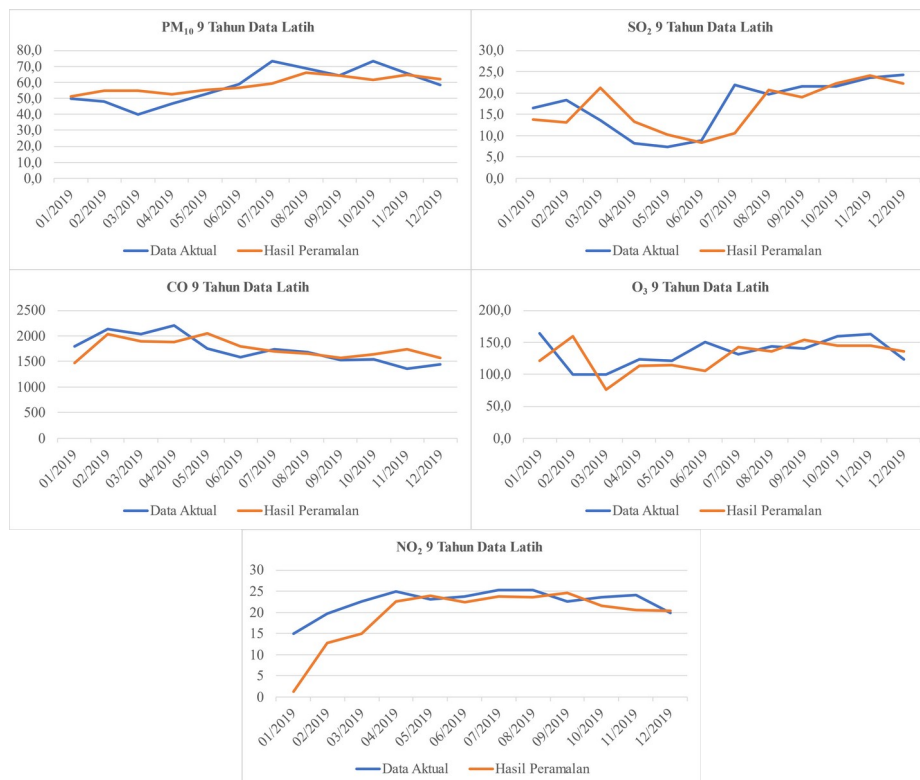
Tabel 6. merupakan hasil terbaik dari Metode SVR dalam peramalan di setiap parameter. Hasil-hasil tersebut jika dirata-ratakan, maka rata-rata MAPE yang didapat adalah 16,19% dan rata-rata RMSE yang didapat adalah 0,1666. Hasil akhir peramalan dengan Metode LSTM setelah proses denormalisasi dapat dilihat pada Tabel 7.

Tabel 7. Hasil Akhir Peramalan SVR

Bu- lan	Hasil Peramalan									
	PM <sub>10</sub> Data Ak- tual	PM <sub>10</sub> Hasil Pera- malan	SO <sub>2</sub> Data Ak- tual	SO <sub>2</sub> Hasil Pera- malan	CO Data Ak- tual	CO Hasil Pera- malan	O <sub>3</sub> Data Ak- tual	O <sub>3</sub> Hasil Pera- malan	NO <sub>2</sub> Data Ak- tual	NO <sub>2</sub> Hasil Pera- malan
Jan- uari 2019	49,8	51,279 56388	16,5	13,809 71465	1800	1465,2 89353	163, 5	120,9 14423 2	15	1,3055 74454
Febr uari 2019	47,9	54,577 95062	18,4	13,036 56561	2133 ,3	2032,4 86104	99,2	159,8 22556 3	19,6	12,778 56924
Mare t 2019	39,8	54,839 51117	13,6	21,197 81894	2037 ,3	1898,7 09854	99,8	76,18 17937 1	22,5	14,889 32522
April 2019	46,5	52,638 45126	8,2	13,198 72215	2200	1878,0 45303	123, 3	113,3 55415 4	24,9	22,552 39381
Mei 2019	52,5	55,309 40834	7,4	10,160 72435	1751 ,4	2049,9 35334	120, 6	114,3 06597 7	23,1	23,877 95432
Juni 2019	58,9	56,590 28918	8,8	8,3184 61441	1582 ,9	1793,6 10276	149, 9	105,4 48753 9	23,8	22,377 85543
Juli 2019	73,2	59,385 0811	21,8	10,518 74536	1741 ,5	1699,3 32391	131, 7	142,4 90355	25,3	23,819 8095

								2		
Agus tus 2019	68,8	65,854 55269	19,7	20,693 63303	1680	1659,4 57734	143, 1	135,5 70514 9	25,3	23,629 79742
Sep- tem- ber 2019	64,3	64,323 15682	21,6	18,932 29487	1523 ,2	1570,8 11005	140, 4	154,1 02325 7	22,5	24,569 53287
Okto- ber 2019	73,4	61,388 02206	21,6	22,291 94475	1540	1640,2 4138	159, 5	144,3 21076 6	23,6	21,589 09479
No- vem- ber 2019	65,6	64,720 79427	23,5	24,167 59562	1363 ,2	1733,9 7507	162, 7	144,7 08251 9	24	20,479 77906
De- sem- ber 2019	58,3	61,858 4141	24,3	22,220 10063	1440	1575,5 72313	123, 3	135,7 88187 6	19,9	20,319 34515

Tabel 7. merupakan tabel perbandingan antara data aktual dari parameter-parameter yang diramalkan dengan hasil peramalan menggunakan Metode SVR. Peramalan yang dilakukan adalah peramalan selama 12 bulan ke depan pada tahun 2019. Visualisasi hasil dapat dilihat pada Gambar 6.



Gambar 6. Visualisasi Hasil Peramalan SVR

Gambar 6. merupakan visualisasi hasil peramalan parameter dengan Metode SVR. Grafik biru merupakan data aktual, sedangkan grafik oranye merupakan data hasil peramalan.

## 5. Kesimpulan

Metode LSTM merupakan metode yang lebih tepat untuk meramalkan kualitas udara di Kota Jakarta Pusat dibandingkan SVR berdasarkan analisis yang dilakukan. Hasil rata-rata tingkat kesalahan dengan *metric* MAPE dan RMSE pada peramalan rata-rata PM<sub>10</sub>, SO<sub>2</sub>, CO, O<sub>3</sub>, dan NO<sub>2</sub> di Kota Jakarta Pusat, menunjukkan bahwa rata-rata kesalahan LSTM lebih kecil dibandingkan SVR. Rata-rata MAPE dari LSTM sebesar 12,15% dan rata-rata RMSE sebesar 0,0941, sedangkan peramalan menggunakan Metode SVR memiliki rata-rata MAPE sebesar 16,19% dan rata-rata RMSE sebesar 0,1666.

#### Daftar Pustaka

- [1] BPS Kota Jakarta Pusat, "Kota Jakarta Pusat Dalam Angka 2022," Jakarta Pusat, 2022. Accessed: Nov. 03, 2022. [Online]. Available: <https://jakpuskota.bps.go.id/publication/download.html?nrbyfeve=YzYwOTY1MWU0NjU2OGQzNDA5MGU0MDIy&xzmn=aHR0cHM6Ly9qYWtwdXNrb3RhLmJwcy5nby5pZC9wdWJsaWNhdGlubi8yMDIyLzAyLzI1L2M2MDk2NTFINDY1NjhhMzQwOTBINDAyMi9rb3RhLWpha2FydGEtcHVzYXQtZGFsYW0tYW5na2EtMjAyMi5odG1s&twoadfnorfeauf=MjAyMi0xMS0wMyAyMDowNT01NA%3D%3D>
- [2] Ismiyati, D. Marlita, and D. Saidah, "Pencemaran Udara Akibat Emisi Gas Buang Kendaraan Bermotor," *Jurnal Manajemen Transportasi & Logistik (JMTransLog)*, vol. 01, no. 03, pp. 241–248, 2014.
- [3] A. H. R. Inaku and C. Novianus, "Pengaruh Pencemaran Udara PM 2,5 dan PM 10 Terhadap Keluhan Pernapasan Anak di Ruang Terbuka Anak di DKI Jakarta," *ARKESMAS (Arsip Kesehatan Masyarakat)*, vol. 5, no. 2, pp. 9–16, 2020, doi: 10.22236/arkemas.v5i2.4990.
- [4] G. E. Bayu, I. K. G. Darma Putra, and N. K. D. Rusjyanthi, "A Comparison Between Backpropagation, Holt-Winter, and Polynomial Regression Methods in Forecasting Dog Bites Cases in Bali," *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, vol. 9, no. 3, p. 251, 2021, doi: 10.24843/jim.2021.v09.i03.p06.
- [5] T. Xayasouk, H. M. Lee, and G. Lee, "Air pollution prediction using long short-term memory (LSTM) and deep autoencoder (DAE) models," *Sustainability (Switzerland)*, vol. 12, no. 6, 2020, doi: 10.3390/su12062570.
- [6] F. Hamami and I. A. Dahlan, "Univariate Time Series Data Forecasting of Air Pollution using LSTM Neural Network," *2020 International Conference on Advancement in Data Science, E-Learning and Information Systems, ICADEIS 2020*, pp. 12–16, 2020, doi: 10.1109/ICADEIS49811.2020.9277393.
- [7] M. C. Turner *et al.*, "Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations," *CA Cancer J Clin*, vol. 70, no. 6, pp. 460–479, 2020, doi: 10.3322/caac.21632.
- [8] M. A. al Ahad, F. Sullivan, U. Demšar, M. Melhem, and H. Kulu, "The effect of air-pollution and weather exposure on mortality and hospital admission and implications for further research: A systematic scoping review," *PLoS One*, vol. 15, no. 10 October, pp. 1–24, 2020, doi: 10.1371/journal.pone.0241415.
- [9] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, and E. Bezirtzoglou, "Environmental and Health Impacts of Air Pollution: A Review," *Front Public Health*, vol. 8, no. February, pp. 1–13, 2020, doi: 10.3389/fpubh.2020.00014.
- [10] A. A. A. Purnamaswari, I. K. G. D. Putra, and I. M. S. Putra, "Komparasi Metode Neural Network Backpropagation dan Support Vector Machines dalam Prediksi Volume Sampah TPA Suwung," *JITTER Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 1, pp. 853–861, 2022, [Online]. Available: <https://ojs.unud.ac.id/index.php/jitter/article/view/83024/43066>
- [11] M. Madhiarasan and S. N. Deepa, "A novel criterion to select hidden neuron numbers in improved back propagation networks for wind speed forecasting," *Applied Intelligence*, vol. 44, no. 4, pp. 878–893, 2016, doi: 10.1007/s10489-015-0737-z.
- [12] H. S. Hota, R. Handa, and A. K. Shrivastava, "Time Series Data Prediction Using Sliding Window Based RBF Neural Network," *International Journal of Computational Intelligence Research*, vol. 13, no. 5, pp. 1145–1156, 2017, [Online]. Available: <http://www.ripublication.com>
- [13] P. Githa Pratiwi, I. Ketut Gede Darma Putra, and D. Purnami Singgih Putri, "Peramalan Jumlah Tersangka Penyalahgunaan Narkoba Menggunakan Metode Multilayer Perceptron," *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, vol. 7, no. 2, p. 143, 2019, doi: 10.24843/jim.2019.v07.i02.p06.

- [14] M. T. Siregar, G. Made, A. Sasmita, G. Agung, and A. Putri, "Perbandingan Analisis Metode Peramalan Jumlah Kasus Penyakit Menular di Kota Bandung (Studi Kasus: Dinas Kesehatan Kota Bandung)," 2022.
  - [15] A. Arfan and L. ETP, "Perbandingan Algoritma Long Short-Term Memory dengan SVR Pada Prediksi Harga Saham di Indonesia," *PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 13, no. 1, pp. 33–43, 2020, doi: 10.33322/petir.v13i1.858.
  - [16] M. F. Rohmah, K. G. D. Putra, R. S. Hartati, and L. Ardiantoro, "Minimum wage correlation with consumer price index predictions using support vector regression," *International Journal of Scientific and Technology Research*, vol. 9, no. 3, pp. 4596–4602, 2020.
-