

# MODEL PREDIKSI TEMPORAL KONSENTRASI POLUTAN UDARA DI PROVINSI DKI JAKARTA MENGGUNAKAN METODE *HYBRID* *ES-RNN* DAN PEMODELAN PREDIKSI DERET WAKTU

ALFREDO Rianto



PROGRAM MAGISTER ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026



## PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA\*

Dengan ini saya menyatakan bahwa tesis dengan judul “Model Prediksi *Temporal* Konsentrasi Polutan Udara di Provinsi DKI Jakarta Menggunakan Metode *Hybrid ES-RNN* dan Pemodelan Prediksi Deret Waktu” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

*Alfredo Rianto*  
G6501221011

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## RINGKASAN

ALFREDO RIAN TO. Model Prediksi Temporal Konsentrasi Polutan Udara di Provinsi DKI Jakarta Menggunakan Metode *Hybrid ES-RNN* dan Pemodelan Prediksi Deret Waktu. Dibimbing oleh IMAS SUKAESIH SITANGGANG dan MUSHTHOFA.

Penelitian ini membahas pengembangan model prediksi *temporal* konsentrasi polutan udara di Provinsi DKI Jakarta menggunakan metode *Hybrid ES-RNN* dan beberapa algoritma *Time Series Forecasting*. Latar belakang penelitian ini didasarkan pada memburuknya kualitas udara Jakarta akibat tingginya aktivitas transportasi, industri, rumah tangga, konstruksi, dan aktivitas perkotaan lainnya. Kondisi tersebut menuntut adanya sistem prediksi yang mampu memperkirakan konsentrasi polutan udara secara lebih akurat agar pemerintah dan pihak terkait dapat melakukan langkah mitigasi lebih cepat. Rumusan masalah penelitian ini adalah bagaimana metode *Hybrid ES-RNN*, LSTM, GRU, dan Facebook Prophet dapat digunakan untuk memprediksi konsentrasi polutan udara terutama PM2.5 dalam rentang waktu Januari 2022 hingga Januari 2023, serta membandingkan model *Hybrid ES-RNN* dengan model *Time Series Forecasting* lainnya. Tujuan penelitian adalah membangun dan mengevaluasi model prediksi konsentrasi polutan udara berdasarkan data historis dan data meteorologi, kemudian membandingkan performa masing-masing algoritma menggunakan metrik evaluasi seperti MAE, MSE, RMSE,  $R^2$ , dan MAPE.

Data penelitian berasal dari data kualitas udara Jakarta dan data meteorologi, antara lain PM10, PM2.5, SO<sub>2</sub>, NO<sub>2</sub>, CO, O<sub>3</sub>, suhu, kelembaban, curah hujan, dan kecepatan angin, namun pada penelitian ini data polutan yang digunakan adalah PM2.5. Data diproses melalui beberapa tahapan, yaitu penggabungan dataset, seleksi fitur, pembersihan data, deteksi *outlier*, penanganan *missing value*, normalisasi, pembagian data latih-validasi-uji, *windowing*, pemodelan, dan evaluasi. Beberapa metode imputasi digunakan, seperti *filling*, interpolasi, *Weighted Moving Average*, *Kalman Filter*, dan kombinasi semua metode. Pada pendekatan *univariat*, model terbaik diperoleh dari *Vanilla LSTM* dengan metode interpolasi, dengan nilai MAE sebesar 0,66, MSE sebesar 0,83, RMSE sebesar 0,91,  $R^2$  sebesar 0,997, dan MAPE sebesar 1,40%. Hasil ini menunjukkan bahwa pola historis PM2.5 sudah cukup kuat untuk memprediksi konsentrasi PM2.5 berikutnya dengan tingkat kesalahan yang rendah. Pada pendekatan *multivariat*, model terbaik diperoleh dari *Stacked LSTM* pada skenario semua metode, dengan nilai MAE sebesar 2,22, MSE sebesar 8,81, RMSE sebesar 2,97,  $R^2$  sebesar 0,97, dan MAPE sebesar 4,44%. Hasil tersebut menunjukkan bahwa *Stacked LSTM* mampu memanfaatkan beberapa variabel *input* untuk menangkap pola perubahan PM2.5, meskipun tingkat kesalahannya masih lebih tinggi dibandingkan pendekatan *univariat* terbaik. Manfaat penelitian ini adalah menyediakan model prediksi kualitas udara yang dapat mendukung pemantauan, peringatan dini, serta pengambilan kebijakan pengendalian pencemaran udara di DKI Jakarta.

Kata kunci: DKI Jakarta, Prediksi Temporal, Polutan Udara, *Stacked LSTM*, *Time Series Forecasting*



## SUMMARY

This study discusses the development of a temporal prediction model for air pollutant concentrations in DKI Jakarta Province using the Hybrid ES-RNN method and several Time Series Forecasting algorithms. The background of this research is based on the declining air quality in Jakarta due to high levels of transportation, industrial, household, construction, and other urban activities. Therefore, a prediction system is needed to estimate air pollutant concentrations more accurately so that the government and related stakeholders can take mitigation measures more quickly. The research problem focuses on how Hybrid ES-RNN, LSTM, GRU, and Facebook Prophet methods can be used to predict air pollutant concentrations, particularly PM<sub>2.5</sub>, within the period from January 2022 to January 2023, as well as to compare the Hybrid ES-RNN model with other Time Series Forecasting models. The objective of this study is to develop and evaluate air pollutant concentration prediction models based on historical and meteorological data, and then compare the performance of each algorithm using evaluation metrics such as MAE, MSE, RMSE,  $R^2$ , and MAPE.

The research data were obtained from Jakarta air quality data and meteorological data, including PM<sub>10</sub>, PM<sub>2.5</sub>, SO<sub>2</sub>, CO, O<sub>3</sub>, NO<sub>2</sub>, temperature, humidity, rainfall, and wind speed, however in this study used PM<sub>2.5</sub> as the pollutant data. The data were processed through several stages, namely dataset merging, feature selection, data cleaning, outlier detection, missing value handling, normalization, train-validation-test splitting, windowing, modeling, and evaluation. Several imputation methods were used, including filling, interpolation, Weighted Moving Average, Kalman Filter, and a combination of all methods. In the univariate approach, the best-performing model was obtained using Vanilla LSTM with the interpolation method, achieving an MAE of 0.66, an MSE of 0.83, an RMSE of 0.91, an  $R^2$  of 0.997, and a MAPE of 1.40%. These results indicate that the historical pattern of PM<sub>2.5</sub> was sufficiently strong to predict subsequent PM<sub>2.5</sub> concentrations with a low error rate. In the *multivariat* approach, the best-performing model was Stacked LSTM under the all-methods scenario, with an MAE of 2.22, an MSE of 8.81, an RMSE of 2.97, an  $R^2$  of 0.97, and a MAPE of 4.44%. These results show that Stacked LSTM was able to utilize several input variables to capture the changing patterns of PM<sub>2.5</sub>, although its error rate was still higher than that of the best univariate approach. The benefit of this study is to provide an air quality prediction model that can support monitoring, early warning systems, and policymaking for air pollution control in DKI Jakarta.

**Keywords:** Air Pollutants, DKI Jakarta, Stacked LSTM, Temporal Prediction, Time Series Forecasting



@Hak cipta milik IPB University

IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
    - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
  2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## © Hak Cipta milik IPB, tahun 2026 Hak Cipta dilindungi Undang-Undang

*Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.*

*Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.*

# **MODEL PREDIKSI TEMPORAL KONSENTRASI POLUTAN UDARA DI PROVINSI DKI JAKARTA MENGGUNAKAN METODE *HYBRID ES-RNN* DAN PEMODELAN PREDIKSI DERET WAKTU**

**ALFREDO RIAN TO**

Tesis

Sebagai salah satu syarat untuk memperoleh gelar

Magister pada

Program Studi Ilmu Komputer

**PROGRAM MAGISTER ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026**





Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji Luar Komisi Pembimbing pada Ujian Tesis:  
Dr. Aziz Kustiyo, S.Si, M.Kom



Judul Tesis : Model Prediksi *Temporal* Konsentrasi Polutan Udara di Provinsi DKI Jakarta Menggunakan Metode *Hybrid ES-RNN* dan Pemodelan Prediksi Deret Waktu

Nama : Alfredo Rianto

NIM : G6501221011

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:

Prof. Dr. Imas Sukaesih Sitanggang, S.Si, M.Kom

---

Pembimbing 2:

Dr. Mushthofa, S.Kom., M.Sc.

---

Diketahui oleh

Ketua Program Studi:

Irman Hermadi, S.Kom., M.S., Ph.D.

NIP 19750311 200604 1 009

---

Dekan Sekolah Sains Data, Matematika, dan Informatika:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.

NIP. 19660702 199302 1 001

---

Tanggal Ujian:  
23 Juni 2026

Tanggal Lulus:



## PRAKATA

Puji dan syukur penulis panjatkan ke hadirat Allah Subhānahu wa Ta'ālā atas segala rahmat, karunia, dan pertolongan-Nya sehingga karya ilmiah ini dapat diselesaikan dengan baik. Tema yang diangkat dalam penelitian yang dilaksanakan sejak bulan Februari 2024 hingga bulan April 2026 ini adalah Model Prediksi *Temporal* Konsentrasi Polutan Udara di Provinsi DKI Jakarta Menggunakan Metode *Hybrid ES-RNN* dan Pemodelan Prediksi Deret Waktu.

Ucapan terima kasih penulis sampaikan kepada dosen pembimbing Prof. Dr. Imas Sukaesih Sitanggang, S.Si M.Kom dan Dr. Mushthofa, S.Kom., M.Sc yang telah memberikan arahan, bimbingan, dan motivasi selama proses penelitian hingga penulisan laporan ini. Ucapan terima kasih juga disampaikan kepada, Bapak Dr. Aziz Kustiyo S.Si., M.Kom atas saran dan masukannya selaku dosen penguji sidang akhir tesis penulis. Penghargaan yang tulus penulis tujukan kepada pihak-pihak yang telah membantu secara langsung maupun tidak langsung, terutama kepada kedua Orang Tua kami Bapak Mauris Adriansyah dan Ibu Rosmaini yang selalu memberikan support dan doa terbaik, keluarga besar, dan pimpinan di kantor, baik dalam bentuk dukungan moral, materi, maupun fasilitas, sehingga penelitian ini dapat terlaksana dengan baik.

Semoga karya ilmiah ini bermanfaat bagi pihak-pihak yang membutuhkan, memberikan kontribusi bagi kemajuan ilmu pengetahuan, serta menjadi amal jariyah yang diridhai oleh Allah Subhānahu wa Ta'ālā.

Bogor, Juli 2026

*Alfredo Rianto*



## DAFTAR ISI

|                                                      |      |
|------------------------------------------------------|------|
| DAFTAR ISI                                           | viii |
| DAFTAR TABEL                                         | ix   |
| DAFTAR GAMBAR                                        | ix   |
| DAFTAR LAMPIRAN                                      | x    |
| PENDAHULUAN                                          | 1    |
| 1.1 Latar Belakang                                   | 1    |
| 1.2 Rumusan Masalah                                  | 3    |
| 1.3 Tujuan                                           | 4    |
| 1.4 Manfaat                                          | 4    |
| 1.5 Ruang Lingkup                                    | 4    |
| II TINJAUAN PUSTAKA                                  | 5    |
| 2.1 Polusi Udara                                     | 5    |
| 2.2 Data Polusi                                      | 7    |
| 2.3 <i>Hybrid Method of ES-RNN</i>                   | 8    |
| 2.4 <i>Long Short-Term Memory</i>                    | 9    |
| 2.5 <i>Stacked LSTM</i>                              | 11   |
| 2.6 Facebook 's Prophet                              | 12   |
| 2.7 <i>Gated Recurrent Unit (GRU)</i>                | 12   |
| 2.8 Evaluasi Model                                   | 13   |
| III METODE                                           | 15   |
| 3.1 Data Penelitian                                  | 15   |
| 3.2 Tahapan Penelitian                               | 18   |
| 3.3 Lingkungan Pengembangan                          | 23   |
| IV HASIL DAN PEMBAHASAN                              | 24   |
| 4.1 Analisis dan Pembersihan Dataset                 | 24   |
| 4.2 Deteksi dan Penanganan <i>Outlier</i>            | 25   |
| 4.3 Proses Penanganan <i>Missing value</i>           | 25   |
| 4.4 Data Polutan Hasil <i>Pra-proses</i>             | 27   |
| 4.5 Normalisasi Dataset                              | 30   |
| 4.6 Pembagian Data Latih, Data Validasi dan Data Uji | 30   |
| 4.7 <i>Windowing (Timestep)</i>                      | 31   |
| 4.8 Pemodelan dan Pelatihan Model                    | 31   |
| 4.9 Evaluasi Model                                   | 32   |
| 4.10 Analisa Model                                   | 36   |
| V SIMPULAN DAN SARAN                                 | 47   |
| 5.1 Simpulan                                         | 47   |
| 5.2 Saran                                            | 48   |
| DAFTAR PUSTAKA                                       | 50   |
| LAMPIRAN                                             | 53   |
| RIWAYAT HIDUP                                        | 59   |

## DAFTAR TABEL

|   |                                                                                                                                                                           |    |
|---|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1 | Jumlah data harian pencemaran polutan di DKI Jakarta tahun 2022 (Dinas Lingkungan Hidup DKI Jakarta 2023)                                                                 | 5  |
| 2 | Deskripsi data <i>multivariat</i>                                                                                                                                         | 7  |
| 3 | Deskripsi atribut meteorologi dari NASA POWER API ( <i>NASA POWER Parameter Dictionary</i> )                                                                              | 15 |
| 4 | Deskripsi Atribut Dataset Kualitas Udara Jakarta Rendah Emisi (Peraturan Menteri Lingkungan Hidup Dan Kehutanan Republik Indonesia Tentang Indeks Standar Pencemar Udara) | 16 |
| 5 | Informasi keterangan kategori (Peraturan Menteri Lingkungan Hidup Dan Kehutanan Republik Indonesia Tentang Indeks Standar Pencemar Udara)                                 | 17 |
| 6 | Statistik deskriptif dataset                                                                                                                                              | 25 |
| 7 | Arsitektur model <i>Hybrid ES-RNN</i> , <i>Vanilla LSTM</i> , <i>Stacked LSTM</i> , <i>Bidirectional LSTM</i> , Model 1 (Haikal 2023), Model LSTM Basic, Model GRU        | 31 |
| 8 | Ringkasan model terbaik pada setiap skenario prediksi PM2.5                                                                                                               | 33 |

## DAFTAR GAMBAR

|    |                                                                                                        |    |
|----|--------------------------------------------------------------------------------------------------------|----|
| 1  | Perbandingan konsentrasi PM2.5 bulanan setiap tahun. Sumber: (Dinas Lingkungan Hidup DKI Jakarta 2023) | 6  |
| 2  | Hubungan <i>neural network</i>                                                                         | 8  |
| 3  | <i>Dataflow and system architecture</i> model <i>ES-RNN</i> (Smyl 2020)                                | 8  |
| 4  | Arsitektur LSTM (Okut 2021)                                                                            | 9  |
| 5  | Struktur <i>Stacked LSTM</i> (Ma <i>et al.</i> 2022)                                                   | 11 |
| 6  | Arsitektur GRU (Gao <i>et al.</i> 2019)                                                                | 13 |
| 7  | Tahapan penelitian prediksi PM2.5 pada pendekatan <i>Univariat</i> dan <i>Multivariat</i>              | 19 |
| 8  | Diagram/ <i>tree</i> skenario eksperimen pada pendekatan <i>univariat</i> dan <i>multivariat</i>       | 20 |
| 9  | Pola perubahan konsentrasi PM2.5 hasil pra-proses menggunakan metode <i>Filling</i>                    | 28 |
| 10 | Pola perubahan kecepatan angin hasil pra-proses menggunakan metode <i>filling</i>                      | 28 |
| 11 | Pola perubahan kelembaban udara hasil pra-proses menggunakan metode <i>filling</i>                     | 29 |
| 12 | Distribusi PM2.5 berdasarkan batas persentil ke-95                                                     | 30 |
| 13 | Data <i>array</i> normalisasi                                                                          | 30 |
| 14 | Data latih, validasi, uji                                                                              | 31 |
| 15 | Plot Ringkasan model terbaik pada setiap skenario prediksi PM2.5                                       | 33 |
| 16 | Hasil prediksi model Prophet terhadap data aktual PM2.5 pada data uji                                  | 37 |
| 17 | Perbandingan nilai aktual dan prediksi PM2.5 Model <i>Hybrid ES-RNN</i> pada data uji                  | 40 |
| 18 | Perbandingan nilai aktual dan prediksi PM2.5 model <i>Stacked LSTM</i>                                 | 42 |
| 19 | Perbandingan nilai aktual dan prediksi PM2.5 model <i>Vanilla LSTM</i>                                 | 43 |
| 20 | Perbandingan nilai aktual dan prediksi PM2.5 model GRU                                                 | 46 |



## DAFTAR LAMPIRAN

|   |                                                                     |    |
|---|---------------------------------------------------------------------|----|
| 1 | Statistik deskriptif dataset hasil penanganan <i>missing value</i>  | 54 |
| 2 | Hasil evaluasi kinerja metode <i>filling</i> terhadap polutan PM2.5 | 55 |
| 3 | Hasil evaluasi kinerja metode WMA terhadap polutan PM2.5            | 56 |
| 4 | Hasil evaluasi kinerja metode Kalman terhadap polutan PM2.5         | 57 |
| 5 | Hasil evaluasi kinerja metode interpolasi terhadap polutan PM2.5    | 57 |
| 6 | Hasil evaluasi kinerja gabungan semua metode terhadap polutan PM2.5 | 58 |