



SKRIPSI

PENERAPAN MODEL CEEMDAN-LSTM DENGAN OPTIMASI BAYESIAN DALAM PREDIKSI INDEKS STANDAR PENCEMAR UDARA DI DKI JAKARTA

RAFIE ISHAQ MAULANA

NPM 21081010073

DOSEN PEMBIMBING

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

Achmad Junaidi, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

PENERAPAN MODEL CEEMDAN-LSTM DENGAN OPTIMASI BAYESIAN DALAM PREDIKSI INDEKS STANDAR PENCEMARAN UDARA DI DKI JAKARTA

RAFIE ISHAQ MAULANA
NPM 21081010073

DOSEN PEMBIMBING

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.
Achmad Junaidi, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

LEMBAR PENGESAHAN

PENERAPAN MODEL CEEMDAN-LSTM DENGAN OPTIMASI BAYESIAN DALAM PREDIKSI INDEKS STANDAR PENCEMAR UDARA DI DKI JAKARTA

Oleh
Rafie Ishaq Maulana
NPM. 21081010073

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi
Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran
Jawa Timur pada tanggal 1 Desember 2025.

Menyetujui

Muhammad Muharrom Al Haromainy,
S.Kom., M.Kom.

NIP. 19950601 202203 1 006



(Pembimbing I)

Achmad Junaidi, S.Kom., M.Kom.

NIP. 19781110 202521 1 048



(Pembimbing II)

Eva Yulia Puspaningrum, S.Kom., M.Kom.

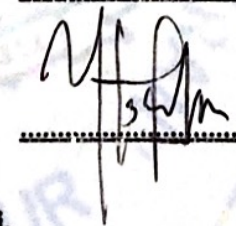
NIP. 19890705 202121 2 002



(Ketua Penguji)

Yisti Vita Via, S.ST., M.Kom.

NIP. 19860425 202121 2 001



(Anggota Penguji)

Mengetahui,

Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**PENERAPAN MODEL CEEMDAN-LSTM DENGAN OPTIMASI
BAYESIAN DALAM PREDIKSI INDEKS STANDAR PENCEMAR
UDARA DI DKI JAKARTA**

Oleh

Rafie Ishaq Maulana

NPM. 21081010073

Menyetujui,

Koordinator Program Studi Informatika

Fakultas Ilmu Komputer



Fetty Tri Anggraeny, S.Kom., M.Kom.

NIP. 19820211 202121 2 005

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama mahasiswa : Rafie Ishaq Maulana
NPM : 21081010073
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu Lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar Pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 1 Desember 2025

Mahasiswa,



Rafie Ishaq Maulana
NPM. 21081010073

ABSTRAK

Nama Mahasiswa / NPM : Rafie Ishaq Maulana / 21081010073

Judul Skripsi : Penerapan Model CEEMDAN-LSTM dengan Optimasi Bayesian Dalam Prediksi Indeks Standar Pencemar Udara di DKI Jakarta

Dosen Pembimbing : 1. Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.
2. Achmad Junaidi, S.Kom., M.Kom.

Kualitas udara yang buruk dapat berdampak negatif terhadap kesehatan pernapasan, sehingga diperlukan sistem prediksi sebagai upaya mitigasi risiko. Namun, data Indeks Standar Pencemar Udara (ISPU) bersifat sangat fluktuatif, sehingga sulit untuk diprediksi secara akurat. Untuk mengatasi permasalahan tersebut, digunakan metode dekomposisi CEEMDAN dan model LSTM. Pemilihan hyperparameter dilakukan menggunakan teknik optimasi Bayesian. Hasil penelitian menunjukkan bahwa kombinasi metode CEEMDAN dan LSTM menghasilkan nilai RMSE sebesar 13.84, MAE sebesar 10.71, dan MAPE sebesar 12.02%. Nilai tersebut menunjukkan peningkatan performa dibandingkan model baseline LSTM tanpa dekomposisi dan optimasi. Dengan skenario pembagian dataset 80:20 dan penggunaan optimizer Adam, model memperoleh hasil evaluasi terbaik, yang menandakan keseimbangan optimal antara kemampuan belajar dan generalisasi. Optimizer Adam terbukti efektif dalam mempercepat konvergensi serta menstabilkan proses pelatihan pada dataset yang digunakan. Penelitian ini juga menunjukkan bahwa penggunaan optimasi Bayesian dalam penentuan hyperparameter tidak memberikan peningkatan performa yang signifikan dibandingkan dengan pemilihan hyperparameter secara manual. Secara keseluruhan, model CEEMDAN-LSTM terbukti mampu meningkatkan akurasi prediksi ISPU, sehingga berpotensi menjadi alat bantu yang efektif dalam mitigasi risiko akibat buruknya kualitas udara.

Kata kunci: ISPU, LSTM, Dekomposisi, CEEMDAN

ABSTRACT

Student Name / NPM : Rafie Ishaq Maulana / 21081010073
Thesis Title : *Application of the CEEMDAN–LSTM Model with Bayesian Optimization for Predicting the Air Pollution Standard Index in DKI Jakarta*
Advisors : 1. Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.
2. Achmad Junaidi, S.Kom., M.Kom.

Poor air quality can have a negative impact on respiratory health, making the development of a prediction system essential as a risk mitigation effort. However, the Air Pollution Standard Index (ISPU) data tends to be highly fluctuating, making it difficult to predict accurately. To address this issue, the CEEMDAN decomposition method and the LSTM model were employed. Hyperparameter selection was carried out using the Bayesian optimization technique. The research results show that the combination of CEEMDAN and LSTM achieved an RMSE of 13.84, MAE of 10.71, and MAPE of 12.02%. These results indicate an improvement in performance compared to the baseline LSTM model without decomposition and optimization. With an 80:20 dataset split and the use of the Adam optimizer, the model achieved the best evaluation results, indicating an optimal balance between learning capability and generalization. The Adam optimizer proved effective in accelerating convergence and stabilizing the training process on the given dataset. The study also revealed that the use of Bayesian optimization for hyperparameter tuning did not yield a significant performance improvement compared to manual hyperparameter selection. Overall, the CEEMDAN-LSTM model proved effective in improving ISPU prediction accuracy, making it a promising tool for mitigating risks associated with poor air quality.

Keywords: *ISPU, LSTM, Decomposition, CEEMDAN*

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Penerapan Model CEEMDAN-LSTM dengan Optimasi Bayesian Dalam Prediksi Indeks Standar Pencemar Udara di DKI Jakarta”** dapat terselesaikan dengan baik. Dalam proses pengerjaan Skripsi penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Muhammad Muharrom Al Haromainy, S.Kom., M.Kom selaku pembimbing pertama yang telah membimbing selama proses pengerjaan skripsi.
4. Bapak Achmad Junaidi, S.Kom., M.Kom selaku pembimbing kedua yang telah membantu dan membimbing selama proses pengerjaan skripsi.
5. Bapak Andreas Nugroho Sihananto, S.Kom., M.Kom selaku PIC Skripsi Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
6. Kedua orang tua, keluarga, dan teman-teman yang selalu memberikan dukungan moral dan semangat hingga skripsi ini dapat terselesaikan.

Penulis Menyadari bahwa laporan skripsi ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kesempurnaan skripsi ini. Akhir kata, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat memberikan manfaat bagi semua pihak.

Surabaya, 24 November 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN.....	iii
LEMBAR PERSETUJUAN.....	iv
SURAT PERNYATAAN BEBAS PLAGIASI.....	v
ABSTRAK.....	vi
ABSTRACT.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	3
1.5. Batasan Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Penelitian Terdahulu.....	5
2.2. Zat Polutan Udara.....	8
2.3. Indeks Standar Pencemar Udara.....	9
2.4. Interpolasi Linear.....	10
2.5. Min-Max Scaler.....	11
2.6. CEEMDAN.....	11
2.7. Sample Entropy.....	13
2.8. K-means.....	14
2.9. LSTM.....	14
2.8.1. Forget Gate.....	15
2.8.2. Input Gate.....	15
2.8.3. Pembaruan Cell State.....	16
2.8.4. Output Gate.....	17
2.10. Fungsi Aktivasi.....	18
2.10.1. Fungsi Sigmoid.....	18

2.10.2. Fungsi Tanh.....	18
2.11. Optimizer.....	19
2.12. Adam Optimizer.....	19
2.13. Optimasi Bayesian.....	19
2.14. Model Surrogate Gaussian Process.....	20
2.15. Fungsi Akuisisi Expected Improvement.....	21
2.16. Pengukuran Nilai Error.....	22
2.17. <i>Root Mean Square Error</i> (RMSE).....	22
2.18. <i>Mean Absolute Error</i> (MAE).....	23
2.19. <i>Mean Absolute Percentage Error</i> (MAPE).....	23
BAB III METODOLOGI.....	24
3.1. Tahapan Penelitian.....	24
3.2. Studi Literatur.....	25
3.3. Desain Pipeline.....	25
3.4. Pengumpulan Data.....	26
3.5. Preprocessing.....	27
3.5.1. Filter Data.....	27
3.5.2. Imputasi Data dengan Interpolasi Linear.....	28
3.5.3. Dekomposisi dengan CEEMDAN.....	29
3.5.4. Hitung Sample Entropy.....	31
3.5.5. Rekonstruksi dengan K-means.....	32
3.5.6. Normalisasi dengan <i>MinMaxScaler</i>	33
3.5.7. Split Data.....	33
3.5.8. Windowing dan Reshape Data.....	34
3.6. Model LSTM.....	35
3.7. Optimasi Bayesian.....	36
3.8. Pembuatan Model.....	38
3.8.1. Set Layer LSTM.....	38
3.8.2. Set Hyperparameter Bound.....	39
3.8.3. Set Initial Points.....	39
3.8.4. Pelatihan Model.....	39
3.9. Evaluasi Model.....	40

3.10. Skenario Pengujian.....	40
3.11. Analisis Hasil.....	41
BAB IV HASIL DAN PEMBAHASAN.....	42
4.1. Pengumpulan Dataset.....	42
4.2. Pemanggilan Dataset.....	42
4.3. Preprocessing.....	44
4.3.1. Filter Data.....	44
4.3.2. Imputasi Data.....	44
4.3.3. Dekomposisi Data.....	45
4.3.4. Hitung Sample Entropy Data.....	47
4.3.5. Klustering dan Rekonstruksi Data.....	49
4.3.6. Split Data.....	50
4.3.7. Normalisasi Data.....	52
4.3.8. Windowing Data.....	54
4.4. Hyperparameter Tuning.....	55
4.4.1. Pembangunan Bayesian Optimization.....	55
4.4.1.1. Pembuatan Fungsi Objektif.....	55
4.4.1.2. Penentuan Search Space.....	56
4.4.2. Inisialisasi Bayesian Optimization.....	57
4.5. Pembangunan Model LSTM.....	59
4.5.1. Arsitektur Model.....	60
4.5.2. Kompilasi Model.....	61
4.6. Pelatihan Model.....	61
4.7. Skenario Pengujian.....	63
4.8. Konfigurasi Model Terbaik.....	65
4.9. Evaluasi Model.....	66
4.9.1. Analisis Grafik.....	66
4.9.2. Evaluasi dengan Metrik Evaluasi.....	68
4.10. Implementasi Model.....	69
BAB V PENUTUP.....	70
5.1. Kesimpulan.....	70
5.2. Saran.....	70

DAFTAR PUSTAKA.....	72
LAMPIRAN.....	75

DAFTAR GAMBAR

Gambar 2.1 Arsitektur LSTM.....	14
Gambar 3.1 Tahapan Penelitian.....	24
Gambar 3.2 Pipeline Model.....	25
Gambar 3.3 Alur Praproses Data.....	27
Gambar 3.4 Alur Proses CEEMDAN.....	29
Gambar 3.5 Alur Rekonstruksi Data.....	31
Gambar 3.6 Alur Model LSTM.....	35
Gambar 3.7 Alur Optimasi Bayesian.....	36
Gambar 3.8 Alur Pembuatan Model.....	38
Gambar 4.1 Dataset ISPU dari Tahun 2010 hingga 2023.....	42
Gambar 4.2 Tampilan Beberapa Baris Pertama Dataset.....	43
Gambar 4.3 Grafik Dataset.....	43
Gambar 4.4 Hasil Filter dan Informasi Dataset.....	44
Gambar 4.5 Data Sebelum dan Sesudah Interpolasi.....	45
Gambar 4.6 Data Hasil Dekomposisi.....	46
Gambar 4.7 Grafik Hasil Dekomposisi.....	47
Gambar 4.8 Sample Entropy Tiap IMF.....	48
Gambar 4.9 Hasil Perhitungan Sample Entropy.....	48
Gambar 4.10 Hasil Rekonstruksi IMF.....	50
Gambar 4.11 Grafik Hasil Rekonstruksi.....	50
Gambar 4.12 Visualisasi Split Data ISPU.....	51
Gambar 4.13 Visualisasi Split Data Co-IMF.....	52
Gambar 4.14 Hasil Normalisasi.....	53
Gambar 4.15 Hasil Windowing.....	54
Gambar 4.16 Hasil Bayesian 1 Layer.....	57
Gambar 4.17 Hasil Bayesian 2 Layer.....	58
Gambar 4.18 Hasil Bayesian 3 Layer.....	59
Gambar 4.19 Ringkasan Model.....	60
Gambar 4.20 Hasil Training Model.....	62
Gambar 4.21 Grafik Hasil Pelatihan Model.....	63

Gambar 4.22 Hasil Pengujian.....	64
Gambar 4.23 Training dan Validation Loss.....	66
Gambar 4.24 Grafik Perbandingan Aktual dan Prediksi.....	67
Gambar 4.25 Hasil Prediksi.....	69

DAFTAR TABEL

Tabel 2.1 Konversi Indeks Standar Pencemaran Udara.....	9
Tabel 2.2 Kategori Rentang Angka ISPU.....	10
Tabel 3.1 Contoh Dataset.....	26
Tabel 3.2 Keterangan Variabel.....	26
Tabel 3.3 Contoh Hasil Filter Dataset.....	27
Tabel 3.4 Contoh Nilai Kosong.....	28
Tabel 3.5 Contoh Hasil Interpolasi Linear.....	28
Tabel 3.6 Contoh Hasil Penambahan White Noise.....	29
Tabel 3.7 Contoh Hasil Perhitungan EMD Ensemble Pertama.....	30
Tabel 3.8 Contoh Hasil IMF1 CEEMDAN.....	30
Tabel 3.9 Contoh Hasil Residual.....	31
Tabel 3.10 Contoh Nilai Sample Entropy.....	32
Tabel 3.11 Perhitungan Jarak ke Centroid.....	32
Tabel 3.12 Hasil Rata-rata Tiap Kluster.....	33
Tabel 3.13 Perhitungan Jarak ke Centroid Baru.....	33
Tabel 3.14 Skema Pembagian Dataset.....	34
Tabel 3.15 Skema Timestep.....	34
Tabel 3.16 Initial Point.....	36
Tabel 3.17 Skenario Tuning Layer LSTM.....	39
Tabel 3.18 Batasan Tuning Hyperparameter.....	39
Tabel 3.19 Pembandingan Model.....	40
Tabel 3.20 Skenario Pengujian.....	40
Tabel 4.1 Hasil Terbaik Bayesian Optimization.....	59
Tabel 4.2 Kombinasi Terbaik.....	66