



UTILIZATION OF HUBERREGRESSOR MACHINE LEARNING MODEL TO PREDICT CARBON MONOXIDE CONCENTRATION IN SURABAYA CITY

Cahya Sugiarto¹⁾, Febby Debora Abigael^{2)*}, Yusron Faiz Athallah³⁾, & Agung Hari Saputra⁴⁾

Sekolah Tinggi Meteorologi Klimatologi dan Geofisika¹⁻⁴
Koresponden*, Email: febbydeboraabigael@gmail.com

Abstract

Carbon monoxide (CO) is one of the pollutant gases whose concentration currently continues to increase due to an increase in population and population activities, especially those that occur in the city of Surabaya, East Java. The purpose of this study is to make a prediction of CO gas concentration in Surabaya City in 2022. CO concentration air quality data was obtained from MERRA-2 Reanalysis through NASA's Giovanni platform. CO concentration data processing is carried out by Machine Learning methods using the Google Colaboratory platform with the HuberRegressor model. The results of the data processing carried out were obtained with details of MASE worth 0.6218, RMSSE worth 0.3657, MAE worth 0.0280, RMSE worth 0.0314, MAPE worth 0.0836, and SMAPE worth 0.0876. From the results of the evaluation of the model, it can be concluded that the HuberRegressor model can make a prediction of CO gas concentration in the city of Surabaya quite well.

Keywords: CO Prediction; HuberRegressor; Machine Learning; Surabaya City

Abstrak

Karbon monoksida (CO) adalah salah satu gas polutan yang saat ini konsentrasinya terus meningkat akibat peningkatan populasi dan aktivitas penduduk terutama yang terjadi di Kota Surabaya, Jawa Timur. Tujuan dari penelitian ini adalah untuk membuat prediksi konsentrasi gas CO di Kota Surabaya tahun 2022. Data kualitas udara konsentrasi CO diperoleh dari MERRA-2 Reanalysis melalui platform Giovanni milik NASA. Pemrosesan data konsentrasi CO dilakukan dengan metode *Machine Learning* menggunakan platform *Google Colaboratory* dengan model HuberRegressor. Hasil dari pemrosesan data yang dilakukan diperoleh hasil dengan rincian MASE senilai 0.6218, RMSSE senilai 0.3657, MAE senilai 0.0280, RMSE senilai 0.0314, MAPE senilai 0.0836, dan SMAPE senilai 0.0876. Dari hasil evaluasi model tersebut, dapat dikatakan bahwa model HuberRegressor dapat membuat prediksi konsentrasi gas CO di Kota Surabaya dengan cukup baik.

Kata Kunci: Prediksi CO; HuberRegressor; Machine Learning; Kota Surabaya

PENDAHULUAN

Seiring dengan berkembangnya zaman, segala aktivitas manusia semakin bergantung pada teknologi. Dengan kemajuan teknologi yang cepat, manusia menjadi lebih mudah mencapai hasil dan

tujuan yang telah ditentukan (Fitri dkk, 2018). Salah satu bukti kemajuan teknologi adalah dengan adanya *Machine Learning* (ML). Pycaret sebagai library untuk ML menyediakan fasilitas penggunaan model HuberRegressor dengan cara yang mudah dan efisien.

HuberRegressor adalah jenis pemodelan regresi yang dirancang untuk mengurangi dampak outlier pada data pelatihan yang tahan terhadap data anomali (Scikit-learn, n.d.). Dalam memprediksi, HuberRegressor mampu menyederhanakan proses prediksi dan memastikan keakuratan hasil prediksi dalam lingkungan data yang bervariasi. Dengan kemampuan pada HuberRegressor, masalah yang sering dijumpai seperti konsentrasi gas polutan berbahaya dapat diprediksi (Rahman, n.d.). Salah satu gas polutan berbahaya yang umumnya dijumpai adalah gas karbon monoksida (CO). Gas CO merupakan gas yang tidak berwarna, tidak berbau, dan tidak berasa (Megalina, 2015). Hadirnya gas CO sangat membahayakan kehidupan manusia bahkan dapat mengakibatkan kematian. Hal ini dikarenakan gas CO mampu membentuk ikatan kuat dengan pigmen darah hemoglobin sehingga berkurangnya penyediaan oksigen ke seluruh tubuh dan menyebabkan sesak nafas kemudian kematian (Sinaga dkk, 2013). Gas CO di udara bersumber dari kendaraan bermotor atau mesin yang menggunakan bahan bakar fosil (Sengkey dkk, 2011). Sehubungan dengan sumber gas CO, daerah dengan penduduk yang padat memiliki potensi penggunaan kendaraan bermotor yang lebih banyak, hal ini berpotensi pada konsentrasi gas CO yang tinggi. Kota Surabaya merupakan kota terbesar kedua di Indonesia setelah Jakarta, sehingga memiliki jumlah penduduk yang sangat banyak dan juga sebagai pusat aktivitas industri (Akhmad dkk, 2017). Berdasarkan Badan Pusat Statistik Surabaya tahun 2021, Kota Surabaya memiliki jumlah penduduk yang mencapai 3 juta jiwa (Nisa & Susanti, 2023). Seiring meningkatnya kepadatan penduduk setiap tahunnya, jumlah kendaraan bermotor yang beredar di masyarakat semakin meningkat, hal ini dapat menyebabkan peningkatan gas polutan termasuk gas CO di udara (Faishol, 2020). Dengan adanya faktor-faktor yang mendukung tingginya

konsentrasi gas CO di Kota Surabaya, diperlukan penanganan dan pencegahan agar bahaya dari gas CO tidak berimbas terhadap Masyarakat. Namun sebelum itu, Langkah awal yang perlu dilakukan adalah memprediksi konsentrasi gas CO di Kota Surabaya. Berdasarkan permasalahan ini, maka dilakukan prediksi konsentrasi gas CO di Kota Surabaya dengan memanfaatkan pemodelan HuberRegressor. Hal ini bertujuan untuk mengetahui sifat konsentrasi CO di Kota Surabaya. Dengan adanya informasi tentang konsentrasi CO, maka akan membantu langkah-langkah apa saja yang perlu dilakukan untuk mencegah dan mengurangi konsentrasi gas CO di Kota Surabaya baik oleh pemerintah maupun warga setempat. Sehingga dampak negatif dari konsentrasi gas CO di Kota Surabaya tidak meluas dan tidak mempengaruhi kehidupan setempat.

METODE

Teknik pengumpulan data

1. Data

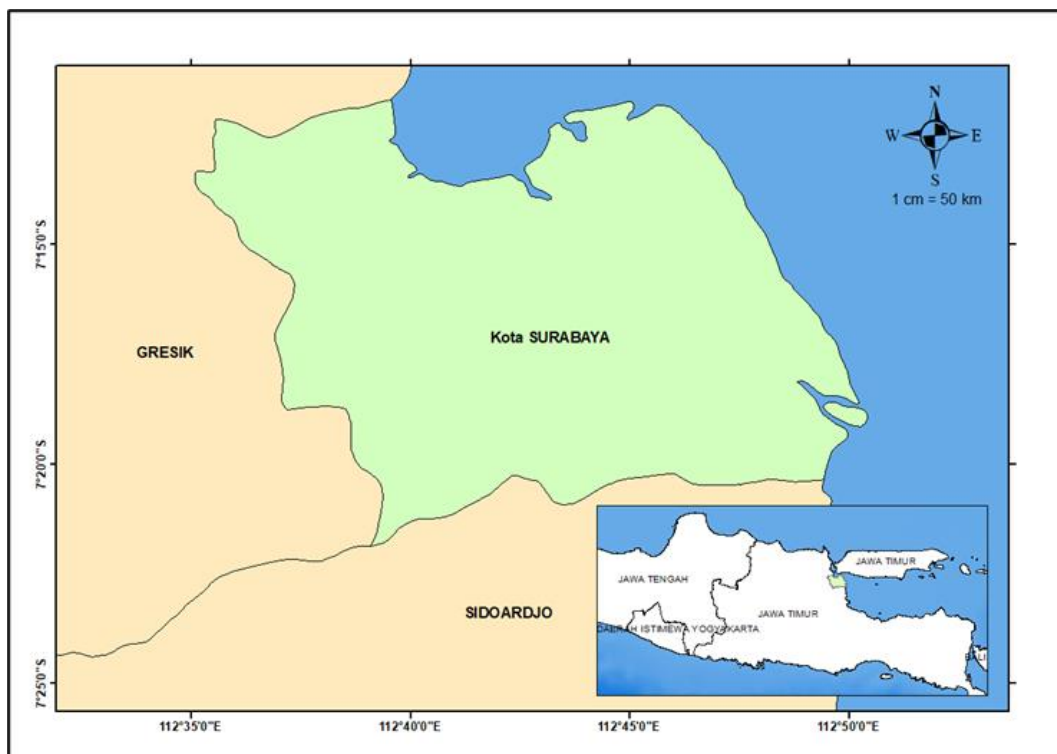
Data yang digunakan adalah salah satu parameter kualitas udara yaitu karbon monoksida (CO). data ini diperoleh dari software Giovanni dengan mengakses web <https://giovanni.gsfc.nasa.gov/giovanni/>. Giovanni (*The Goddard Interactive Online Visualization and Analysis Infrastructure*) merupakan *software* yang memberikan akses eksplorasi, analisis, dan pengunduhan data atmosfer yang dihasilkan oleh berbagai satelit dan instrumen penginderaan jauh milik NASA (Leptoukh, 2007). Data yang digunakan adalah data MERRA-2 *Reanalysis* (*The Modern-era Retrospective Analysis for Research and Applications Version 2*) yang merupakan data analisis gabungan antara data observasi dengan data model (Alfiandy & Permana, 2020). Data ini merupakan

data bulanan dengan rentang waktu mulai tahun 2013 hingga tahun 2022. Pemilihan rentang waktu ini dikarenakan data sepuluh tahun dapat memberikan gambaran yang luas dalam memprediksi data.

2. Lokasi

Lokasi yang digunakan adalah Kota Surabaya. Kota Surabaya merupakan ibukota Provinsi Jawa Timur. Letak astronomis Kota Surabaya adalah pada $7^{\circ}9'-7^{\circ}21'$ LS dan $122^{\circ}36'-112^{\circ}54'$ BT dengan luas sekitar 33.048 ha (Rosytha, 2016). Kota Surabaya berbatasan dengan Selat Madura, Kabupaten Sidoarjo, dan Kabupaten

Gresik (BPS Surabaya, 2015). Kota Surabaya dipilih karena merupakan salah satu kota metropolitan setelah Jakarta yang memiliki aktivitas padat di dalamnya (Tajuddin & Trilaksana, 2015). Selain itu, Kabupaten Sidoarjo dan Kabupaten Gresik merupakan daerah yang memiliki aktivitas industri yang tinggi. Aktivitas industri yang tinggi pada kedua kabupaten tersebut akan mempengaruhi konsentrasi CO pada Kota Surabaya sebagai daerah yang berbatasan langsung dengan kedua kabupaten tersebut.



Gambar 1. Peta Kota Surabaya

Metode pengolahan data

Langkah awal yang perlu dilakukan yaitu mengunduh data menggunakan peramban *website* Giovanni. Hal pertama yang dilakukan adalah masuk ke dalam *website* Giovanni dan melakukan login. Setelah berhasil masuk, selanjutnya adalah memilih *Time Series, Recurring Averages* dalam menu *select plot*, memilih *Measurements CO*, menentukan rentang

waktu data bulanan mulai dari bulan Januari 2013 hingga bulan Desember 2022, memilih lokasi yaitu Kota Surabaya, kemudian plot data. Setelah data berhasil di plot, selanjutnya adalah mengunduh data pada menu *download* dengan format CSV. CSV adalah format data di mana setiap data atau *record* di dalamnya dipisahkan menggunakan koma (,) atau titik koma (;) (Rafael, 2020). Tahapan berikutnya, pengolahan dataset dilakukan

menggunakan perangkat lunak excel. Data diimpor ke excel untuk beberapa tahapan, seperti penghapusan data yang tidak diperlukan, pengelompokkan data menjadi dua variabel (tanggal dan Konsentrasi CO), penyederhanaan nilai konsentrasi CO menjadi bentuk desimal dengan dua angka dibelakang koma (dimana data mentah

berupa konsentrasi CO dikalikan 10^{-9} dengan satuan kilogram per meter persegi per detik). Tujuan dari penyederhanaan nilai konsentrasi ini adalah untuk efisiensi komputasi dan mengurangi *overfitting*. Kemudian dataset disimpan di *Google Drive* dalam format file XLSX.

Tabel 1. Lima data terawal dari dataset yang sudah diolah

Tanggal	Konsentrasi CO
01/01/2013	1.58
01/02/2013	1.62
01/03/2013	1.63
01/04/2013	1.60
01/05/2013	1.58

Setelah pengolahan dataset dilakukan, selanjutnya adalah penginstalan library Pandas, PyCaret, dan Mathplotlib menggunakan *Google Colaboratory* (Colab) kemudian menghubungkan *Google Drive* dengan Colab untuk dapat memanggil dataset yang sudah disimpan dan menyimpan hasil prediksi. Dalam hal ini, Pycaret menyediakan *tools* untuk *preprocessing* data, membangun model, membandingkan model, mengevaluasi, menampilkan visualisasi hasil training sampai menyimpan hasil model dalam

format tertentu (Mustaqim, 2020). PyCaret akan menguji banyak model kemudian menggunakan model terbaik untuk melakukan prediksi (Abidin & Nuryana, 2023). Pada Langkah ini, Dataset yang sudah dipanggil ke dalam Colab, kemudian dilakukan konfigurasi *hyperparameter* seperti *data*, *target*, *fh*, *session_id*, *fold*, *verbose*, *remove_harmonics*, *seasonal_period*, *fold_strategy*, *max_sp_to_consider*, *harmonic_order_method*, *num_sps_to_use*, dan *coverage*.

Tabel 2. Konfigurasi hyperparameter

Deskripsi	Nilai
<i>target</i>	Konsentrasi CO
<i>fh</i>	12
<i>session_id</i>	123
<i>fold</i>	5
<i>verbose</i>	<i>True</i>
<i>remove_harmonics</i>	<i>True</i>
<i>seasonal_period</i>	12
<i>fold_strategy</i>	<i>sliding</i>
<i>max_sp_to_consider</i>	48
<i>harmonic_order_method</i>	<i>harmonics_strength</i>
<i>num_sps_to_use</i>	1
<i>coverage</i>	0.7

Konfigurasi *hyperparameter* dilakukan untuk mendapatkan model yang dapat melakukan prediksi dengan baik. Data hasil prediksi model kemudian divisualisasikan dalam bentuk grafik deret waktu. Selanjutnya verifikasi hasil prediksi dilakukan dengan membandingkan data

prediksi model terhadap data observasi, menggunakan koefisien korelasi, MASE, RMSSE, MAE, RMSE, MAPE, dan SMAPE untuk mengetahui tingkat keakuratan model.

Teknik analisis data

Dari data yang telah diproses diperoleh beberapa output berupa analisis dekomposisi yang mencakup data grafik data *seasonal*, *trend*, dan *residual*.

1. Seasonal

Data seasonal merupakan data yang menunjukkan adanya perubahan dengan berpola repetitif yang cukup teratur baik dalam skala harian, bulanan, musiman, maupun tahunan (Rahmadianto, 2022).

$$Seasonal = \sum_{j=1}^k (a_j \cdot \cos(\frac{2\pi jt}{p}) + b_j \cdot \sin(\frac{2\pi jt}{p}))$$

Dengan k adalah istilah Fourier yang digunakan a_j dan b_j adalah koefisien untuk istilah Fourier ke-j, dan t adalah waktu.

2. Trend

Trend merupakan komponen yang menggambarkan pola perubahan jangka panjang dalam data deret waktu baik peningkatan maupun penurunan (Utama, 2007).

$$Trend = T(t_0) + k \cdot \Delta t_t = m + k \cdot (t_1 - t_0)$$

Dengan t_1 dan t_2 adalah waktu 1 dan waktu 2, k adalah laju pertumbuhan, dan m adalah *offset*.

3. Residual

Residual adalah perbedaan selisih dari data yang sebenarnya atau data actual dengan data hasil prediksi yang telah diperoleh (Agustina, 2022).

$$Nilai\ Residual = Nilai\ Actual - Nilai\ Prediksi$$

Analisis Diagnostik Data

Permodelan prediksi konsentrasi CO yang dilakukan juga menghasilkan Analisis Diagnostik berupa fungsi ACF dan PACF.

1. ACF (*Auto Correlation Function*)

ACF adalah suatu fungsi yang digunakan pada data deret statistika terutama dalam permodelan prediksi. Pada ACF dilakukan pengukuran

keterkaitan hubungan atau korelasi dengan suatu variabel dengan bilangannya sendiri pada kondisi waktu yang berbeda (Tando dkk, 2016).

$$\rho_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}$$

Dengan ρ_k adalah koefisien korelasi pada lag - k dan \bar{x} = rata-rata observasi.

2. PACF (*Partial Auto Correlation Function*)

PACF merupakan suatu fungsi yang juga digunakan dalam data deret statistika terkhusus model prediksi. Pada PACF dilakukan pengukuran hubungan keterkaitan suatu variabel dengan bilangannya sendiri pada waktu yang berbeda dan dengan cara menghapus pengaruh dari lag sebelumnya atau waktu yang lalu.

$$\phi_{k+1,k+1} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{kj} \rho_{k+1-j}}{1 - \sum_{j=1}^{k-1} \phi_{kj} \rho_j}$$

Dengan ρ_k adalah koefisien korelasi pada lag - k.

Evaluasi model

Untuk menguji kualitas performa dari suatu model data prediksi dapat dilakukan dengan evaluasi model. Evaluasi model ini mencakup beberapa skala error antara lain: MASE (*Mean Absolute Scaled Error*), RMSSE (*Root Mean Squared Scaled Error*), MAE (*Mean Absolute Error*), RMSE (*Root Mean Squared Error*), MAPE (*Mean Absolute Percentage Error*), dan SMAPE (*Symmetric Mean Absolute Percentage Error*) (Mu'minin dkk, 2022) (Anwar, 2023) (Arif dkk, 2023) (Wiranda & Sadikin 2019) (Aji, n.d.).

$$MASE = \frac{MAE}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}$$

$$RMSSE = \sqrt{\frac{MSE}{\frac{1}{T-1} \sum_{t=2}^T (y_t - y_{t-1})^2}}$$

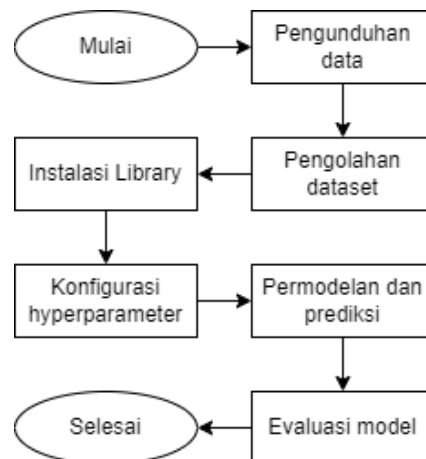
$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}$$

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left(\frac{|y_t - \hat{y}_t|}{y_t} \right) \times 100\%$$

$$SMAPE = \frac{1}{T} \sum_{t=1}^T \frac{2 \times |y_t - \hat{y}_t|}{|y_t| + |\hat{y}_t|} \times 100\%$$

Dengan T adalah jumlah observasi, y_t adalah nilai aktual pada waktu t, dan \hat{y}_t adalah nilai prediksi pada waktu t.

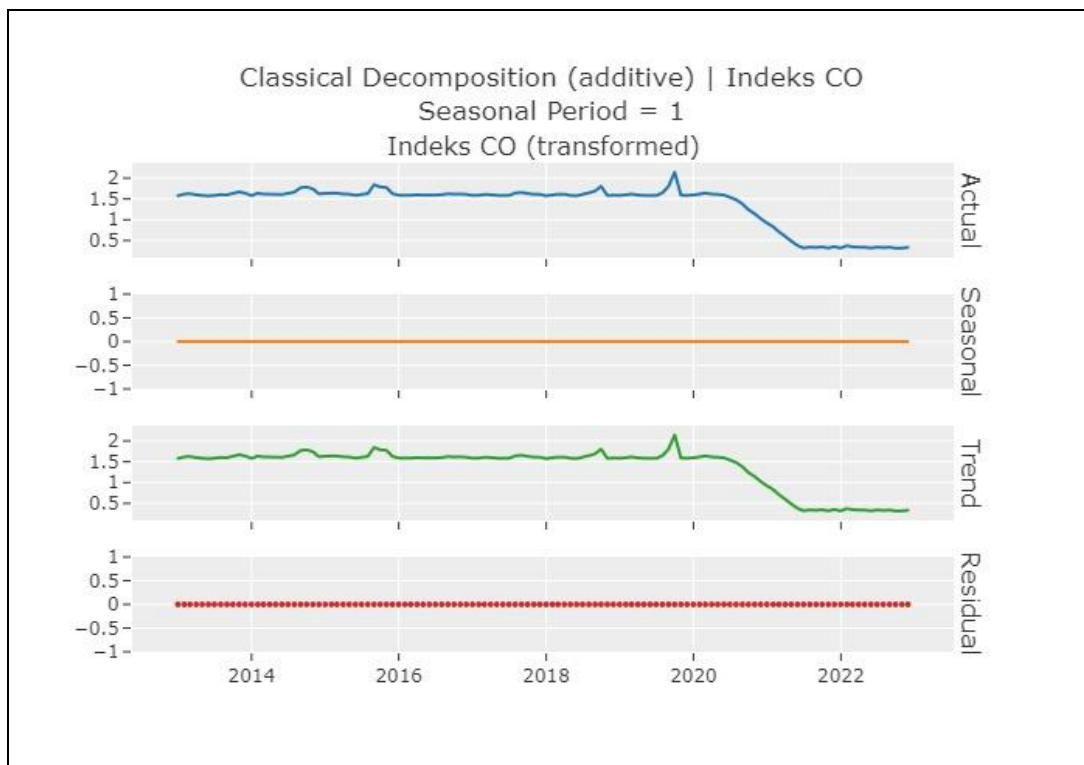


Gambar 2. Diagram alir penelitian

HASIL DAN PEMBAHASAN

Analisis Dekomposisi

Berikut grafik actual, seasonal, trend, dan residual konsentrasi CO di Kota Surabaya.



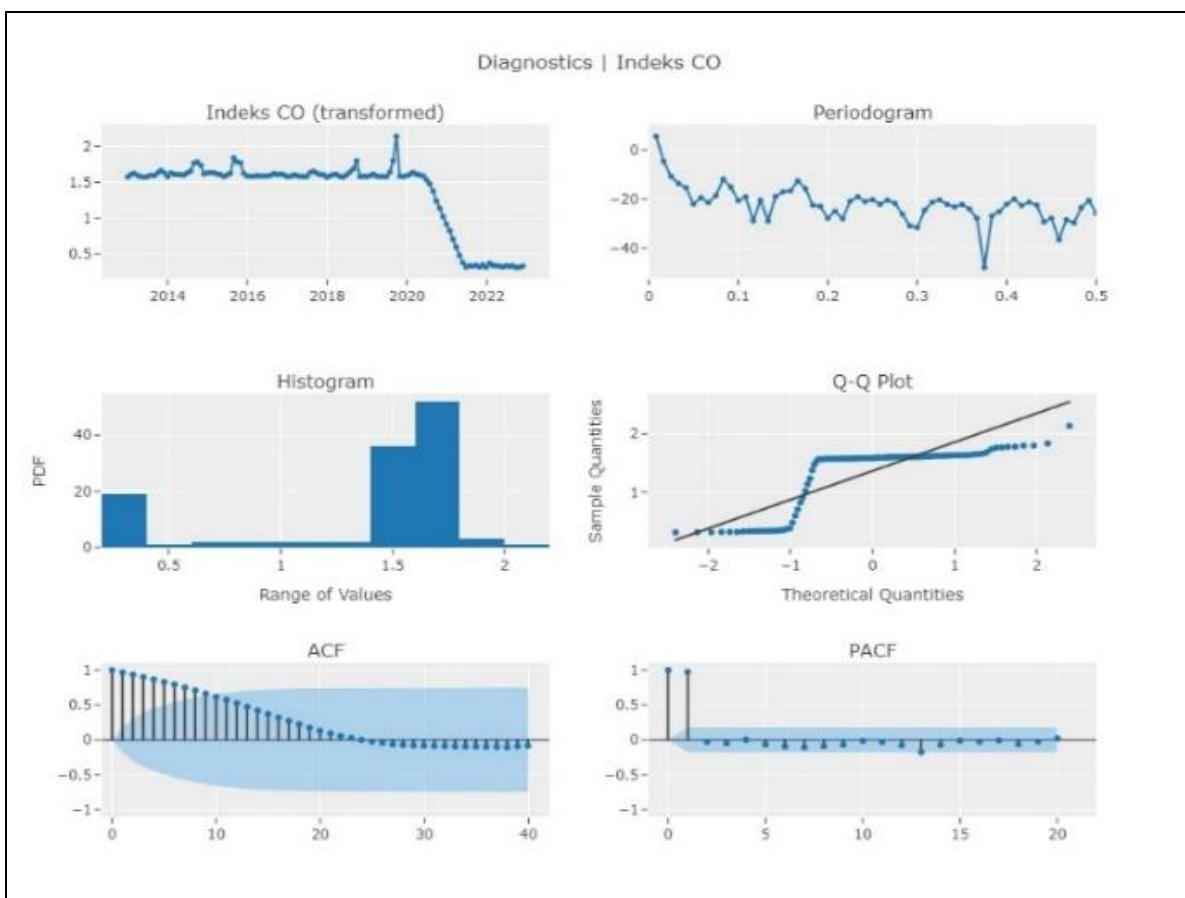
Gambar 3. Actual, Seasonal, Trend, dan Residual

Berdasarkan grafik dekomposisi klasik di atas, komponen *actual* atau keadaan sebenarnya menunjukkan konsentrasi CO pada tahun 2013 hingga 2020 berada pada sekitar 1.5 hingga 2 $\mu\text{gm}^{-2}\text{s}^{-1}$ meskipun terdapat lonjakan hingga melewati 2 $\mu\text{gm}^{-2}\text{s}^{-1}$ pada tahun 2019. Pada tahun 2020, konsentrasi CO menurun hingga di bawah 0.5 $\mu\text{gm}^{-2}\text{s}^{-1}$ dan secara konstan di bawah 0.5 $\mu\text{gm}^{-2}\text{s}^{-1}$ hingga pada tahun 2022. Pada grafik *seasonal* menunjukkan angka 0. Hal ini menunjukkan bahwa tidak ada perubahan yang terlalu signifikan pada skala tahunan. Pada grafik trend, terlihat

bahwa konsentrasi CO di Surabaya pada tahun 2013 hingga 2022 mengalami penurunan khususnya pada tahun 2020. Sedangkan pada grafik residual menunjukkan angka 0. Hal ini menunjukkan bahwa selisih antara nilai actual dengan prediksi tidak terlalu besar atau tidak signifikan.

Analisis diagnostik

Berikut grafik konsentrasi CO (*transformed*), periodogram, histogram, Q-Q Plot, ACF, dan PACF di Surabaya.



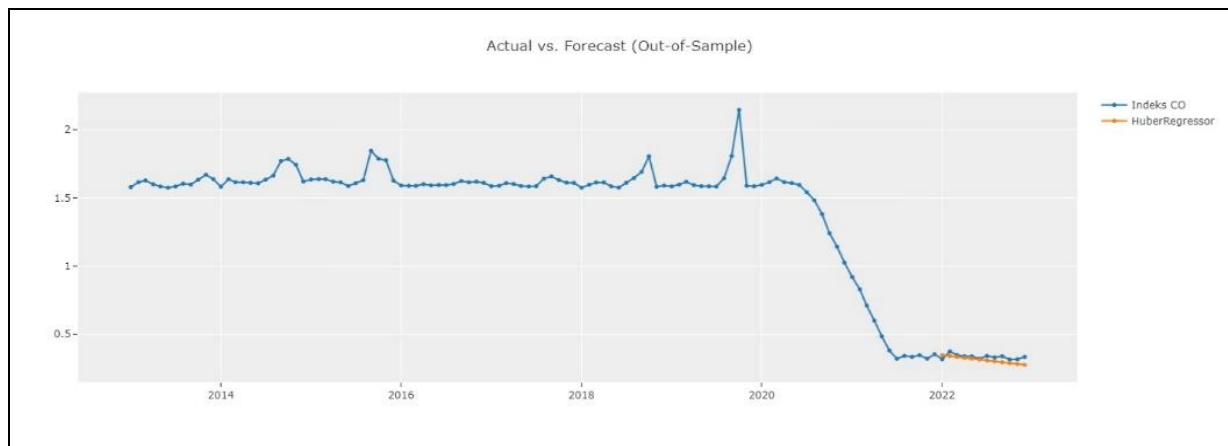
Gambar 4. Grafik konsentrasi CO (*Transformed*), Periodogram, Histogram, Q-Q Plot, ACF, dan PACF

Pada konsentrasi CO (*transformed*), terlihat penurunan tajam di akhir periode yang menandakan terjadinya peristiwa signifikan yang mempengaruhi konsentrasi CO di Kota Surabaya. Penurunan konsentrasi CO terjadi pada sekitar pertengahan tahun 2020 hingga 2021. Pada periodogram, menunjukkan bahwa tidak ada puncak yang menonjol

dimana hal itu menunjukkan tidak ada siklus atau pola musiman yang kuat dalam data. Pada histogram, terlihat bahwa distribusi tidak simetris dengan sebagian besar data berkumpul di nilai yang lebih tinggi. Nilai paling tinggi berada pada rentang nilai diantara 1.5 dan 2 dengan nilai PDF (*Probability Density Function*) diatas 40. Pada Q-Q Plot, terdapat titik-titik

yang meyimpang dari garis lurus yang menunjukkan bahwa data tidak terdistribusi normal. Pada ACF, sumbu x menyatakan lag, sedangkan sumbu y menyatakan koefisien *autocorrelation*. Grafik ACF menunjukkan bahwa pada lag awal, korelasi yang diberikan adalah positif dengan nilai maksimum mendekati 1, sedangkan lag berikutnya menurun dengan nilai minimum kurang dari -0.5 sehingga grafik ACF menunjukkan

terdapat beberapa lag awal yang menunjukkan korelasi signifikan yang menunjukkan bahwa adanya autokorelasi yang positif pada lag tersebut. Bar yang lebih tinggi dan di luar area biru (*confidence band*) menunjukkan bahwa ada korelasi pada lag tertentu yaitu signifikan secara statistik. Sedangkan pada PACF, hanya terdapat satu atau dua lag yang menunjukkan korelasi parsial signifikan.



Gambar 5. Prediksi v.s. Aktual

Pada grafik, *actual* adalah nilai yang sebenarnya terjadi, sedangkan *forecast* adalah nilai yang diprediksi. Pada grafik, nilai *actual* ditunjukkan oleh garis berwarna biru, sedangkan nilai *forecast* ditunjukkan oleh garis berwarna orange dengan menggunakan Huberregressor. Huberregressor sendiri adalah model regresi linear yang tahan terhadap *outlier* atau data yang menyimpang dari *trend* umum data. Hubungan nilai *actual* dan

forecast menunjukkan bahwa prediksi yang dilakukan cukup baik meskipun terdapat penyimpangan namun tidak terlalu signifikan.

Evaluasi model

Berikut tabel evaluasi model yang mencakup MASE, RMSSE, MAE, RMSE, MAPE, SMAPE, dan R2.

Tabel 3. Evaluasi Model

Model	MASE	RMSSE	MAE	RMSE	MAPE	SMAPE
Huberregressor	0.6218	0.3657	0.0280	0.0314	0.0836	0.0876

Tabel di atas menunjukkan bahwa model Huberregressor memiliki nilai di bawah 1 yang menunjukkan bahwa model melakukan prediksi yang lebih baik daripada metode *naive forecast*. Nilai 0.6218 pada MASE menunjukkan bahwa model peramalan memiliki akurasi yang cukup baik. Nilai 0.3657 dan 0.0314 pada RMSSE dan RMSE merupakan nilai yang

rendah, hal ini menunjukkan bahwa kesalahan yang terjadi lebih kecil. Nilai 0.0280 pada MAE menunjukkan rata-rata kesalahan antara prediksi dan nilai sebenarnya hanya sedikit. Nilai 0.0836 dan 0.0876 pada MAPE dan SMAPE menunjukkan bahwa model memiliki akurasi yang lebih baik.

KESIMPULAN

Berdasarkan analisis data yang telah dibahas, diperoleh bahwa terdapat penurunan signifikan nilai konsentrasi CO di Surabaya, tidak menunjukkan pola siklikal atau musiman yang kuat, distribusi nilai tidak normal dengan kecenderungan pada nilai yang lebih tinggi, terdapat autokorelasi signifikan pada lag awal, dan terdapat korelasi parsial yang signifikan pada satu atau dua lag awal. Nilai 0.6218 pada MASE, 0.3657 pada RMSSE, 0.0280 pada MAE, 0.0314 pada RMSE, 0.0836 pada MAPE, dan 0.0876 pada SMAPE menunjukkan bahwa model dapat melakukan kesalahan prediksi yang lebih kecil serta error yang kecil. Hal ini menunjukkan bahwa model Huberregressor dapat digunakan dalam memprediksi konsentrasi CO di Kota Surabaya dengan cukup baik.

DAFTAR PUSTAKA

- Abidin, A. R., & Nuryana, I. K. D. (2023). Perbandingan Metode Klasifikasi Data Mining Untuk Mengukur Tingkat Kepuasan Mahasiswa Terhadap Sistem Informasi Penilaian Nonakademik UNESA (SIPENA). *Journal of Emerging Information System and Business Intelligence (JEISBI)*, 4(4), 129-138.
- Agustina, I. (2022). Analisis Pengembangan Model Prediksi Kesuksesan Kickstarter Menggunakan Algoritma Backpropagation Artificial Neural Network (ANN) dan Random Forest.
- Aji, I. M. S. *Prediksi Data Sensor Kelembapan Tanah Yang Rusak Menggunakan Algoritma K-Means dan Support Vector Regression* (Bachelor's thesis, Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta).
- Akhmad, M. W., Vitianingsih, A. V., & Wijaya, T. A. (2017). Pemetaan Tingkat Polusi Udara di Kota Surabaya Berbasis Android. *Inform: Jurnal Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 1(1).
- Alfiandy, S., & Permana, D. S. (2020). Tren Curah Hujan Berbasis Data Sinoptik BMKG dan Reanalisis MERRA-2 NASA di Provinsi Sulawesi Tengah. *Jurnal Sains & Teknologi Modifikasi Cuaca*, 21(2), 63-72.
- Anwar, G. S. (2023). *Implementasi dan Optimalisasi Metode Single Moving Average untuk Memprediksi Stok Barang Percetakan dan ATK* (Doctoral dissertation, Universitas Siliwangi).
- Arif, Z., Adhi, N. A. S., & Muttaqin, A. (2023). Penerapan Jaringan Syaraf Tiruan dan Certainty Factor untuk Peramalan Penjualan Sepeda Motor. *REMIK: Riset dan E-Jurnal Manajemen Informatika Komputer*, 7(3), 1523-1533.
- BPS Surabaya. (2015, 10 Januari). Surabaya adalah Ibu Kota Provinsi Jawa Timur yang Dikenal sebagai Kota Pahlawan. Diakses pada 27 Desember 2023, dari <https://surabayakota.bps.go.id/statictable/2015/01/10/441/surabaya-adalah-ibu-kota-propinsi-jawa-timur-vang-dikenal-sebagai-kota-pahlawan.html>
- Faishol, M. A. (2020). *Analisis Data Runtun Waktu Prediksi Polusi Udara di Kota Surabaya menggunakan Deep Learning RNN-LSTM* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- Leptoukh, G. (2007, April). Exploring NASA and ESA atmospheric data using Giovanni, the online visualization and analysis tool. In *ENVISAT Symposium*.
- M. C. Mustaqim, "Membangun Model Machine Learning dengan Mudah Menggunakan PyCaret | by MOCH. CHAMDANI MUSTAQIM | Medium," 2020. <https://medium.com/@chamdani/membangun-model-machine-learning-dengan-mudah-menggunakan-pycaret-29f36d2c283c> (diakses 29 Desember 2023).
- Megalina, Y. (2015). Pengaruh pencemaran udara di daerah terminal Amplas bagi kehidupan masyarakat. *Jurnal Pengabdian Kepada Masyarakat*, 21(79), 94-101.
- Mu'minin, F., Fauziah, F., & Gunaryati, A. (2022). Prediksi Kunjungan Wisatawan Mancanegara Melalui Pintu Udara Menggunakan ARIMA, Glmnet, dan Prophet. *Techno. Com*, 21(1), 149-157.
- Nisa, L. C., & Susanti, A. (2023). Strategi Penerapan Mobil Listrik di Surabaya Sebagai Smart Mobility. *Jurnal MITRANS (Media Publikasi Terapan Transportasi)*, 1(2 (Agustus)), 213-225.
- Rafael, R. (2020). *PENERAPAN ALGORITMA PGP UNTUK ENKRIPSI CSV FILE DI PT. X* (Doctoral dissertation, Universitas Atma Jaya Yogyakarta).
- Rahmadianto, B. (2022). Prediksi Harga Saham BBKA dan BMRI dengan Model Seasonal ARIMA.

- Rosytha, A. (2016). Studi dampak pengembangan pemukiman di wilayah pesisir surabaya timur. *Agregat*, 1(2).
- S. Fitri, N. Nurjanah, and W. Astuti, "Penerapan Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa (Studi Kasus: UMTAS)," *Simetris J. Tek. Mesin Elektro Dan Ilmu Komput.*, vol. 9, no. 1, pp. 633–640, Apr. 2018, doi: 10.24176/simet.v9i1.2002.
- Scikit-learn. (n.d.). 1.1. *Linear Models*, diakses pada 11 Januari 2024, dari https://scikit-learn.org/stable/modules/linear_model.html
- Sengkey, S. L., Jansen, F., & Wallah, S. E. (2011). Tingkat pencemaran udara CO akibat lalu lintas dengan model prediksi polusi udara skala mikro. *Jurnal Ilmiah Media Engineering*, 1(2).
- Sinaga, S., Sudarno, S., & Handayani, D. S. (2013). Pengaruh Jumlah Kendaraan Dan Faktor Meteorologi Terhadap Konsentrasi Karbon Monoksida (Co) Di Jalan Pandanaran Kawasan Simpang Lima, Kota Semarang (Doctoral dissertation, Diponegoro University).
- Tajuddin, R., & Trilaksana, A. (2015). Perubahan Tradisi Ritual Sedekah Bumi di Kota Metropolitan Surabaya: Analisa Perubahan Tradisi Ritual Sedekah Bumi di Dusun Jeruk Kelurahan Jeruk Kecamatan Lakarsantri Kota Surabaya Tahun 1990-2014.
- Tando, J., Komalig, H., & Nainggolan, N. (2016). Prediksi Jumlah Penumpang Kapal Laut di Pelabuhan Laut Manado Menggunakan Model ARMA. *d'CARTESIAN: Jurnal Matematika dan Aplikasi*, 5(2), 95-99.
- Utama, N. P. S. (2007). Prakiraan Kebutuhan Tenaga Listrik Propinsi Bali Sampai Tahun 2018 Dengan Metode Regresi Berganda Deret Waktu. *Majalah Ilmiah Teknologi Elektro*, 6(1).
- Wiranda, L., & Sadikin, M. (2019). Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, 8(3), 184-196.
- Y. Rahman, "Penerapan Machine Learning dalam E-Commerce Perusahaan Bukalapak".