



PERAMALAN KONSENTRASI PM2.5 MENGGUNAKAN MODEL ARCH/GARCH DAN LONG SHORT-TERM MEMORY (STUDI KASUS: KOTA JAKARTA PUSAT)

Fardhi Dzakwan Fauzan¹, Dhyas Adhyza Rayhan¹, Hala Mutiara Putri¹, Fitri Kartiasih^{2,*}

¹Program Studi DIV Komputasi Statistik, Politeknik Statistika STIS, Indonesia

²Program Studi DIV Statistika, Politeknik Statistika STIS, Indonesia

Abstrak: Polusi udara merupakan masalah serius di Jakarta akibat revolusi industri dan aktivitas komuter yang tak pernah berhenti. *Particulate matter* (PM) 2.5 merupakan salah satu dari 6 polutan berupa partikel tersuspensi berukuran diameter aerodinamis lebih kecil dari 2,5 μm . PM2.5 menyebabkan terganggunya sistem kardiovaskular dan respiratory, mengakibatkan kejadian kelahiran prematur, dan kejadian berat badan lahir rendah. Oleh karena itu, dibutuhkan kebijakan dan aturan dari pemerintah terkait pengendalian polusi udara agar masyarakat memperoleh kesehatan yang lebih baik, mengurangi kemacetan, mitigasi perubahan iklim, dan efisiensi energi. Peramalan konsentrasi PM2.5 menjadi salah satu hal penting yang dilakukan untuk pengambilan kebijakan demi tercapainya visi jangka panjang tahun 2030, yaitu kota yang aman, nyaman, produktif, berkelanjutan, sebanding dengan kota besar lainnya dan rumah bagi warga (Menuju Udara Bersih Jakarta). Penelitian ini bertujuan untuk mendapatkan model terbaik untuk peramalan konsentrasi PM2.5 di Wilayah Jakarta Pusat. Metode yang digunakan, yaitu *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) dengan *Long Short-Term Memory* (LSTM) dan menguji akurasi model dengan data sebenarnya. Hasil penelitian memberikan kesimpulan bahwa metode LSTM lebih baik dalam meramalkan konsentrasi PM2.5 berdasarkan MAE, MAPE, MSE, dan RMSE. Hasil peramalan dengan metode LSTM menunjukkan bahwa konsentrasi PM2.5 di Wilayah Jakarta Pusat selama 48 jam kedepan berada di rentang tidak sehat bagi kelompok sensitif dan di beberapa jam berikutnya masuk ke kategori tidak sehat. Maka dari itu, dibutuhkan perhatian lebih dari pemerintah untuk melakukan pengendalian polusi udara demi kualitas udara dan kesehatan masyarakat kota.

Kata kunci: ARCH/GARCH, Jakarta Pusat, LSTM, PM2.5

I. PENDAHULUAN

Revolusi industri mengakibatkan kesuksesan besar umat manusia dalam perkembangan teknologi, masyarakat, dan penyediaan layanan. Di sisi lain, aktivitas manusia akibat dari revolusi industri tersebut berdampak buruk terhadap lingkungan dengan mencemari

air, udara, dan tanah sehingga memproduksi sejumlah besar polutan yang berakibat pada kesehatan manusia (Manisalidis et al., 2020). Seperti contoh penggunaan perangkat elektronik yang dapat menyebabkan lebih banyak emisi di udara sehingga meningkatkan pemanasan global (Br Perangin-angin et al., 2024; Elheddad et al., 2021; Pribadi & Kartiasih, 2020). Selain itu, akibat dari perluasan industri untuk meningkatkan kebutuhan energi juga menyebabkan peningkatan emisi polutan (Kartiasih &

*) 222112042@stis.ac.id, fkartiasih@stis.ac.id

Diterima: 9 Februari 2024

Direvisi: 10 Mei 2024

Disetujui: 14 Mei 2024

DOI: 10.23969/infomatek.v26i1.12603

Setiawan, 2020; Shafiei & Salim, 2014). World Health Organization (WHO) juga mengatakan bahwa polusi udara bisa sangat parah dan berdampak pada anak-anak yang tinggal di negara berpenghasilan rendah-menengah yang memiliki tingkat kualitas udara yang buruk (Com, 2018). Anak-anak mempunyai risiko khusus akibat terpapar polusi udara karena organ dan sistem di dalam tubuh mereka masih dalam tahap perkembangan. Selain itu, mereka juga memiliki laju pernapasan yang lebih tinggi dibandingkan orang dewasa sehingga lebih banyak menghirup udara yang tercemar (American Lung Association, 2023). Oleh karena itu, polusi udara antropogenik merupakan salah satu bahaya kesehatan saat ini di dunia karena menyebabkan 9 juta kematian setiap tahunnya (WHO, 2023).

World Health Organization (WHO) menjelaskan bahwa terdapat 6 polutan udara utama, yaitu *particulate matter* (PM), ozon di permukaan tanah, karbon monoksida, sulfur oksida, nitrogen oksida, dan timbal (Jayadri et al., 2024; Manisolidis et al., 2020; Swardanasuta et al., 2024). *Particulate matter* (PM) merupakan sumber racun bagi kesehatan (Mukherjee & Agrawal, 2017). *Particulate Matter* (PM) terbentuk karena adanya reaksi kimia di atmosfer antara berbagai polutan. *Particulate matter* baik PM_{2.5} atau PM₁₀ mengandung tetesan kecil cairan atau padatan yang dapat menyebabkan efek kesehatan yang serius (Cheung et al., 2011). Namun, partikel halus PM_{2.5} memiliki risiko kesehatan yang lebih besar (L. Zhang et al., 2019). Secara khusus, PM_{2.5}, partikel tersuspensi yang diameter aerodinamisnya lebih kecil dari 2,5 µm merupakan polutan partikel yang paling terkenal dan telah menjadi masalah di seluruh dunia (G. Zhang et al., 2018). Berdasarkan WHO Global Air Quality Guidelines (2021), konsentrasi rata-rata

tahunan PM_{2.5} tidak melebihi 5 µg/m³ dan NO₂ tidak melebihi 10 µg/m³ dan pada musim puncak rata-rata konsentrasi O₃ selama 8 jam tidak melebihi 60 µg/m. Fakta lain menunjukkan bahwa partikel PM_{2.5} dapat tertahan di atmosfer dalam waktu yang lama serta menyebar ke tempat lain karena dimensinya yang sangat kecil (W. E. Wilson & Suh, 1997). Partikel halus tersebut salah satu penyebab “kabut” di berbagai daerah metropolitan (Kan et al., 2012). Sumber PM_{2.5} dapat bervariasi sesuai dengan lokasi, kondisi iklim, sumber emisi, dan pola penyebaran (Thangavel et al., 2022). Jika ditinjau dari lokasi, berbagai sumber dapat berkontribusi dalam konsentrasi PM_{2.5} seperti emisi kendaraan, resuspensi debu, garam laut, emisi industri, emisi pesawat terbang, emisi kapal, dan pembakaran biomassa (Manousakas et al., 2017). Komponen utama yang menyusun PM_{2.5}, yaitu karbon hitam (Bové et al., 2019), *polycyclic aromatic hydrocarbons* (Leachi et al., 2020), *aryl hydrocarbons* (Zhao et al., 2019), *volatile organic hydrocarbons* (Puri et al., 2017), logam berat (Manisolidis et al., 2020), senyawa organik (Grzywa-Celińska et al., 2020), mineral (Kitamura et al., 2018), ion organik (Švédová et al., 2020), dan bahan biologis (Shahrbaf et al., 2021).

Penelitian membuktikan bahwa PM_{2.5} dapat memberikan dampak ke sistem yang ada di dalam tubuh. Efek jangka panjang dan jangka pendek terpapar PM_{2.5} mempengaruhi efek kesehatan. Dilihat dari rentang waktu terpapar PM_{2.5}, jangka pendek memengaruhi sistem kardiovaskular yang menyebabkan peningkatan peluang seseorang untuk serangan jantung (Wyatt et al., 2022). Pada sistem respiratory, dapat menyebabkan aritmia, asma, bronkiolitis, dan mengi (Chen et al., 2021; Lei et al., 2019; Pini et al., 2021; Y. Zhang et al., 2020). Di sisi lain, paparan dalam

jangka panjang akan mengakibatkan kejadian kelahiran prematur, dan kejadian berat badan lahir rendah (Xie et al., 2021), peningkatan risiko alzheimer, parkinson, dan penyakit neurodegeneratif (Wang et al., 2021). Ditambah lagi penelitian mengenai gangguan spektrum autisme, menunjukkan bahwa PM2.5 kemungkinan terkait dengan penyakit perkembangan saraf (Costa et al., 2020). Studi lain juga menunjukkan bahwa ada hubungan antara PM2.5 dan risiko kanker paru-paru (Hamra et al., 2014). Hubungan antara penyakit pernapasan yang diakibatkan oleh virus dan polusi udara juga sudah diketahui secara luas. Beberapa penelitian menyebutkan bahwa penyebaran COVID-19 juga dipengaruhi oleh polutan. Cantini et al (2020) memiliki hipotesis bahwa atmosfer dengan konsentrasi polutan yang tinggi, kondisi meteorologi, akan mendorong partikel virus di udara untuk waktu yang lama, sehingga akan mengakibatkan penularan virus COVID-19. Hal ini didukung dengan penelitian yang dilakukan oleh (Martelletti & Martelletti, 2020) bahwa wilayah utara Italia yang paling terkena dampak COVID-19, yaitu wilayah dengan konsentrasi PM10 dan PM2.5 yang besar.

Provinsi DKI Jakarta merupakan salah satu kota metropolitan yang sering disebut sebagai Jakarta Metropolitan Area (JMA). Sebagai kota metropolitan, aktivitas di Provinsi DKI Jakarta tidak pernah berhenti terutama aktivitas komuter. Di wilayah perkotaan, lalu lintas komuter memiliki hubungan yang signifikan dengan PM2.5 yang terbentuk (Chaney et al., 2017). Hal ini mengindikasikan bahwa PM2.5 yang terbentuk di Jakarta merupakan salah satu akibat dari adanya emisi kendaraan komuter. Berdasarkan data IQAir, Jakarta menempati peringkat 12 kota besar paling berpolusi. Beberapa hari sebelumnya, Jakarta menempati posisi ke 8,

bahkan pernah menjadi kota paling berpolusi ke 4 di dunia. Sumber utama polutan utama di DKI Jakarta adalah PM2.5 dengan konsentrasi lebih dari 10x lipat nilai panduan kualitas udara tahunan WHO. Berdasarkan publikasi Statistik Komuter Jabodetabek Badan Pusat Statistik (2019), 11,1% dari 29,3 juta penduduk Jabodetabek berumur 5 tahun ke atas merupakan komuter. Sementara itu, Kota Jakarta Selatan merupakan tujuan terbesar komuter Jabodetabek. Kegiatan utama komuter Jabodetabek, yaitu bekerja (80,5%). Sebesar 72% komuter menggunakan kendaraan pribadinya sebagai transportasi utama dalam kegiatan komuter nya. Sekitar 90% dari mereka tidak memiliki keinginan untuk beralih ke moda transportasi umum dengan alasan waktu tempuh yang lama (36,6%), tidak praktis (34,9%), akses yang jauh (8,4%), dan tidak aman (1,7%). Dalam sebulan terakhir pada saat pencacahan survei komuter tahun 2019, komuter Jabodetabek mengalami keluhan demam (15,1%), batuk/pilek (27,2%), sakit kepala (19,2%), sakit tenggorokan (7,3%), sakit mata (1,7%), sesak nafas (1,7%), masuk angin (31,6%), dan pegal-pegal (37,2%). Diperkuat dengan data dari Dinas Kesehatan (Dinkes) DKI Jakarta, bahwa ditemukan sebanyak 638.291 kasus ISPA di periode Januari hingga Juni 2023. Data IQAir (2023) menyebutkan bahwa akibat dari polusi udara, 8.100 penduduk Jakarta kehilangan nyawanya dan membawa kerugian sebesar Rp32,09 triliun. Penelitian yang dilakukan dari organisasi kesehatan global Vital Strategis dan Dinas Lingkungan Hidup DKI Jakarta mengestimasi bahwa polusi udara yang ada di Jakarta berpotensi menyebabkan lebih dari 10.000 kematian setiap tahun. Selain itu, penduduk Jakarta merupakan penduduk yang kehilangan jumlah tahun hidup tertinggi dibandingkan dengan provinsi lain di Indonesia. Hal ini disebabkan

karena kesehatan yang buruk, disabilitas, dan kematian dini akibat dari paparan PM2.5 (Menuju Udara Bersih Jakarta).

Pemerintah Provinsi DKI Jakarta menyadari bahwa polusi di Jakarta merupakan masalah yang serius. Dinilai bahwa polusi memiliki dampak terhadap beban kesehatan dan ekonomi di Jakarta. Hal ini termasuk biaya kesehatan, ongkos kematian, serta kehilangan produktivitas akibat penyakit dan perawatan penyakit yang disebabkan oleh polusi (G. Zhang et al., 2018). Sampai saat ini, Pemerintah Provinsi DKI Jakarta telah mengeluarkan beberapa kebijakan dan aturan terkait pengendalian polusi udara di sektor transportasi, industri pengolahan, dan pembangkit listrik. Hal ini juga dilakukan agar masyarakat memperoleh kesehatan yang lebih baik, mengurangi kemacetan, mitigasi perubahan iklim, dan efisiensi energi. Pemerintah Provinsi DKI Jakarta melalui publikasi *Menuju Udara Bersih Jakarta* mengemukakan bahwa kualitas udara di Jakarta terus menurun sejauh perkembangan kota itu sendiri, tetapi bukan tidak mungkin perkembangan kota Jakarta disertai dengan peningkatan kualitas udara. Pemerintah Provinsi DKI Jakarta terus menggandeng *stakeholder* terkait untuk merumuskan kebijakan yang kuat dan evaluasi berbasis bukti yang nyata untuk hasil terbaik dalam pengurangan polusi (Menuju Udara Bersih Jakarta). Hal ini juga menjadi salah satu visi jangka panjang tahun 2030 agar menjadi kota yang aman, nyaman, produktif, berkelanjutan, sebanding dengan kota besar lainnya dan rumah bagi warga (Menuju Udara Bersih Jakarta).

Prediksi PM2.5 menjadi salah satu hal penting yang dilakukan untuk pengambilan kebijakan demi tercapainya visi jangka panjang tahun 2030. Konsentrasi PM2.5 dapat diukur pada

suatu titik dengan fasilitas penunjang. Banyak metode yang sudah dilakukan untuk memprediksi konsentrasi PM2.5 pada area tertentu dengan variasi akurasi (G. Zhang et al., 2018). Metode *time series* saat ini masih menjadi salah satu pendekatan yang paling sering digunakan untuk mengestimasi dan memprediksi konsentrasi PM2.5 (G. Zhang et al., 2018). Model *time series* yang digunakan dibagi menjadi 3 kategori, yaitu model tradisional, model *machine learning*, dan model *deep learning* (Shi et al., 2022). Model tradisional kemudian terbagi menjadi model linier, seperti *Autoregressive Moving Average* (ARMA) (G. T. Wilson, 2016), *Autoregressive Integrated Moving Average* (ARIMA) (G. T. Wilson, 2016), dan *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) (Galbraith & Zinde-walsh, 2001), dan model non-linier seperti *Autoregressive Conditional Heteroskedasticity* (ARCH), *Generalized ARCH* (GARCH), *Exponential Generalized ARCH* (EGARCH), *Threshold Autoregressive* (TAR), *Non-linear Autoregressive* (NAR), dan *Non-linear Moving Average* (NMA) (Shi et al., 2022). Sementara itu, perkembangan teknologi metode *machine learning* seperti *Support Vector Machine* (SVM), *Artificial Neural Network* (ANN) (Nor et al., 2017), dan *deep learning NNs* (Masum et al., 2018) terbukti memberikan performa yang lebih baik dibandingkan dengan metode tradisional. Di sisi lain, pendekatan *machine learning* untuk *time series forecasting* yang lebih cocok digunakan mencakup RNN, LSTM (Sahoo et al., 2019), dan GRU.

Penelitian ini mencakup beberapa kontribusi baru, yaitu dengan mencari model terbaik menggunakan metode *deep learning Long Short-Term Memory* (LSTM) dengan metode tradisional ARCH/GARCH. Resolusi peramalan dilakukan per jam, lebih tinggi dibandingkan penelitian terdahulu yang

umumnya hanya per hari sehingga hasil peramalan lebih akurat dan bermanfaat untuk keperluan pengambilan kebijakan operasional pencegahan polusi. Studi kasus di Wilayah Jakarta Pusat dengan karakteristik Jakarta Pusat sebagai tempat kegiatan ekonomi dan transportasi terpadat. Akibatnya, emisi gas buang kendaraan serta potensi peningkatan level polusi dapat terjadi pada jam-jam sibuk di Wilayah Jakarta Pusat. Oleh karena itu, hasil peramalan konsentrasi PM2.5 per jam di Jakarta Pusat dapat memberikan manfaat praktis yang lebih besar bagi berbagai pihak dalam melakukan langkah-langkah pencegahan dan pengendalian peningkatan konsentrasi PM2.5 pada jam-jam rawan. Hasil ramalan PM2.5 ini dapat dimanfaatkan dengan lebih aplikatif oleh pemangku kepentingan dalam mengantisipasi dan mencegah peningkatan polusi per jam di Jakarta Pusat.

II. METODOLOGI

2.1 Data

Data diperoleh dari website iqair.com yang merupakan sebuah website yang menyajikan laporan kualitas udara secara *real time* dan prakiraan polusi udara di seluruh dunia. Secara spesifik, data diperoleh dari stasiun pemantau kualitas udara Jakarta, GBK. Namun, data yang diperoleh terdapat beberapa *missing data* (data yang hilang) sehingga perlu dilakukan pendekatan untuk menanganinya. Data yang hilang dapat mengancam validitas uji coba dan dapat menyebabkan kesimpulan yang tidak valid (Kang, 2013).

Data yang digunakan pada penelitian ini adalah data sekunder periode 27 September Pukul 19.00 WIB hingga 10 November 2023 Pukul 01.00 WIB. Terdapat total 1.041 data dengan 23 data hilang (*missing value*). Nilai median setiap jam diambil sebagai

representasi untuk menghindari pengaruh nilai ekstrem. Penelitian ini hanya menggunakan data PM2.5 dalam rentang 2 bulan. Penelitian selanjutnya disarankan menggunakan dataset PM2.5 dalam rentang waktu setidaknya satu tahun sehingga didapatkan pola data yang lebih bervariasi. Selain itu faktor-faktor yang memengaruhi konsentrasi PM2.5 selain faktor waktu dapat dipertimbangkan sebagai variabel independen.

2.2. Metodologi

2.2.1 Preprocessing

Preprocessing dilakukan untuk menangani data yang hilang. Pada tahap ini, dilakukan interpolasi data menggunakan model LSTM. LSTM dilatih menggunakan data konsentrasi PM2.5 yang lengkap, kemudian digunakan untuk memprediksi atau menginterpolasi nilai-nilai yang hilang. Dengan demikian diharapkan nilai hasil interpolasi mendekati nilai aktualnya. Setelah data dilengkapi, langkah berikutnya yang dilakukan adalah plotting data guna melihat pola pergerakan data *time series* konsentrasi PM2.5. Plotting data penting dilakukan untuk analisis lebih lanjut dalam pemodelan, misalnya untuk menentukan stasioneritas data. Dengan data yang sudah lengkap, analisis selanjutnya dapat dilaksanakan.

2.2.2 ARCH/GARCH

ARCH/GARCH merupakan metode analisis untuk memprediksi volatilitas data *time series*. Metode ini membantu mengatasi masalah heteroskedastisitas bersyarat atau perubahan variasi data dari waktu ke waktu. Pertama, data dibagi menjadi data *training* dan data *testing* sebesar masing-masing 80% dan 20%. Kemudian, data *training* digunakan sebagai acuan dalam pembuatan model untuk melakukan *forecasting*. Data *training* kemudian dilakukan uji stasioneritas dengan melihat pola ACF dan PACF dan uji formal

dengan *unit root test*. Adapun uji *unit root test* yang digunakan yaitu uji *Augmented Dickey-Fuller* yang memiliki persamaan sebagai berikut :

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \sum_{i=2}^n \beta \Delta Y_{t-1} + e_t \quad (1)$$

Dengan melihat uji tersebut, jika ditemukan data tidak stasioner di level, proses *differencing* perlu dilakukan dengan persamaan sebagai berikut

$$\Delta Y_t = Y_t - Y_{t-1} \quad (2)$$

Setelah melihat kestasioneran data, dilanjutkan dengan menerapkan metode *Box-Jenkins* untuk mencari model terbaik dengan menelusuri pola ACF, PACF, maupun percobaan data *training*. Metode *Box-Jenkins* terdiri dari 3 model, yaitu *Autoregressive* (AR), *Moving Average* (MA), *Autoregressive Moving Average* (ARMA), dan model campuran *Autoregressive Integrated Moving Average* (ARIMA). Dengan persamaan model sebagai berikut:

Model *Autoregressive* AR(p) :

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (3)$$

Model *Moving Average* MA(q) :

$$Y_t = \delta + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (4)$$

Model *Autoregressive Moving Average* ARMA(p,q) :

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (5)$$

Model *Autoregressive Integrated Moving Average* (ARIMA) :

$$Y_t = \delta + (1 + \phi_1)Y_{t-1} + \dots + (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (6)$$

Model terbaik dicari dengan mengestimasi parameter-parameter dalam model tersebut. Hasil estimasi parameter kemudian diuji dengan melakukan uji signifikansi parameter dengan cara membandingkan p-value dengan tingkat signifikansi (α) dalam uji hipotesis. Setelah dilakukan uji signifikansi parameter, dilakukan identifikasi efek ARCH dengan menganalisis pola residual kuadrat dengan uji korelogram dan uji formal ARCH-LM untuk melihat heteroskedastisitas dalam data *training*. Efek ARCH ditandai dengan adanya pola berulang pada korelogram dan hasil uji yang signifikan pada uji formal ARCH-LM . Ketika efek ARCH ditemukan, maka proses dilanjutkan dengan mencari kandidat model dan yang akan digunakan. Model varians yang dapat digunakan, yaitu model ARCH dan GARCH dengan persamaan sebagai berikut:

Model *Autoregressive Conditional Heteroscedasticity* ARCH(p)

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 \quad (7)$$

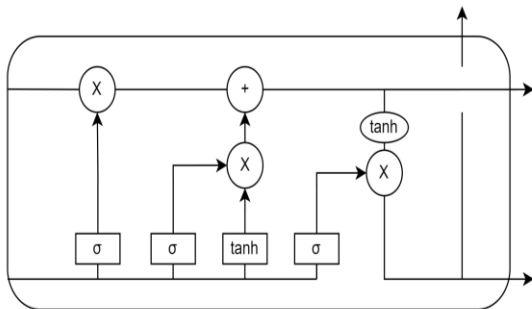
Model *Generalized Autoregressive Conditional Heteroscedasticity* GARCH(p,q)

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \quad (8)$$

Model terbaik yang didapatkan kemudian dievaluasi melalui *goodness of fit* dengan uji statistik AIC (*Akaike Information Criterion*). Kemudian model rata-rata terbaik dan model varians terbaik di gabungkan dengan membentuk model baru untuk melakukan peramalan terhadap konsentrasi PM2.5 di Wilayah Jakarta Pusat. Model baru yang terbentuk kemudian dievaluasi performanya dalam meramalkan konsentrasi PM2.5 di Wilayah Jakarta Pusat dengan menggunakan ukuran uji statistik seperti MSE, RMSE, MAE, dan MAPE.

2.2.3 Long Short-Term Memory (LSTM)

Peneliti menggunakan algoritma LSTM dengan bantuan bahasa pemrograman *python* untuk melakukan *forecasting*. Menurut penelitian yang pernah dilakukan mengatakan bahwa model *deep learning*, yaitu *Artificial Neural Network* (ANN) dan LSTM, memiliki kinerja interpolasi yang lebih baik daripada model konvensional, seperti *Linear Regression* (LR) dan *Support Vector Regression* (SVR) untuk menangani data polusi udara yang hilang di Taiwan (Yen et al., 2020). Penelitian lain juga mengatakan bahwa metode imputasi data yang hilang berbasis LSTM memberikan akurasi prediksi konsentrasi PM2.5 yang lebih baik dibandingkan metode konvensional, seperti *mean-imputation* dan *imputation-Moving Average* (Yuan et al., 2018).



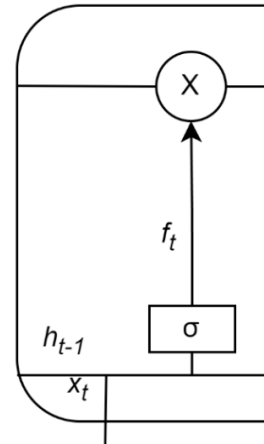
Gambar 1. Arsitektur Long Short Term Memory (LSTM) ([geeksforgeeks.org](https://www.geeksforgeeks.org))

Dalam struktur LSTM, informasi disimpan oleh sel dan manipulasi memori dilakukan oleh *gate*. Terdapat tiga *gate* yaitu *forget gate*, *input gate*, dan *output gate*. *Forget gate* berperan untuk memutuskan jenis informasi yang harus dibuang atau disimpan dari keadaan sel yang diformulasikan dalam bentuk sebagai berikut:

$$f_t = \sigma(W_f S_{t-1} + W_f X_t) \quad (9)$$

Dengan W_f adalah bobot dari *forget gate*; S_{t-1} adalah *state* sebelumnya atau

state pada waktu $t-1$; X_t adalah *input* pada waktu t ; dan σ adalah fungsi aktivasi *sigmoid*.

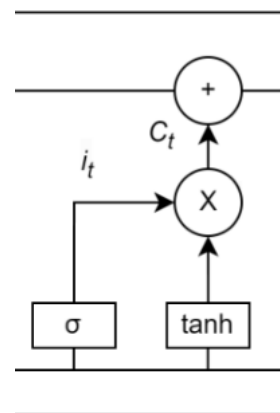


Gambar 2. Forget Gate ([geeksforgeeks.org](https://www.geeksforgeeks.org))

Input gate berperan untuk mengontrol informasi baru apa yang akan ditambahkan ke status sel dari *input* saat ini yang diformulasikan dalam bentuk sebagai berikut:

$$i_t = \sigma(W_i S_{t-1} + W_i X_t) \quad (10)$$

Dengan W_i adalah bobot dari *input gate*; S_{t-1} adalah *state* sebelumnya atau *state* pada waktu $t-1$; X_t adalah *input* pada waktu t ; dan σ adalah fungsi aktivasi *sigmoid*.



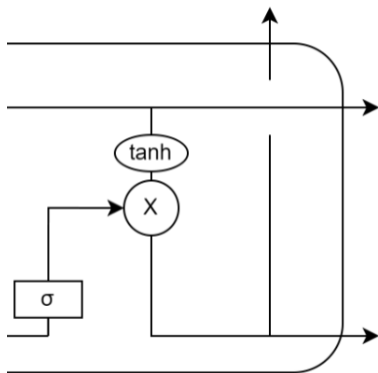
Gambar 3. Input Gate ([geeksforgeeks.org](https://www.geeksforgeeks.org))

Terakhir, *Output gate* berperan untuk menentukan berapa nilai keadaan tersembunyi berikutnya dalam setiap langkah waktu yang diformulasikan dalam bentuk sebagai berikut:

$$o_t = \sigma(W_o S_{t-1} + W_o X_t) \quad (11)$$

$$h_t = o_t * \tanh(\tanh(c_t)) \quad (12)$$

Dengan W_o adalah bobot dari *output gate*; S_{t-1} adalah *state* sebelumnya atau *state* pada waktu $t-1$; X_t adalah *input* pada waktu t ; dan σ adalah fungsi aktivasi *sigmoid*.



Gambar 4. *Output Gate (geeksforgeeks.org)*

Penelitian ini dimulai dengan membangun model LSTM dengan 3 lapisan. Lapisan pertama merupakan lapisan LSTM standar dengan menghasilkan keluaran berupa urutan, bukan hanya waktu terakhir saja. Lalu, lapisan kedua yang dibuat merupakan lapisan LSTM *bidirectional* dimana model memproses data maju atau mundur untuk memperoleh pemahaman yang lebih baik tentang pola data. Kemudian yang terakhir merupakan lapisan LSTM *bidirectional* yang hanya mengeluarkan keluaran pada waktu terakhir saja. Dari model yang dibuat, model kemudian dilatih dengan data *training*. Kemudian di evaluasi *training loss* dan *validation loss* untuk melihat performa model dalam belajar terhadap data *training*. Setelah terlihat indikasi

bahwa model belajar dengan baik terhadap data yang diberikan, model akan digunakan untuk melakukan peramalan konsentrasi PM2.5 di Wilayah Jakarta Pusat. Terakhir, model akan dievaluasi performanya dalam meramalkan konsentrasi PM2.5 di Wilayah Jakarta Pusat dengan menggunakan ukuran uji statistik seperti MSE, RMSE, MAE, dan MAPE.

3.2.4 Evaluasi Model

Mean Square Error (MSE) adalah ukuran kesalahan yang berbasis pada kuadrat perbedaan antara nilai observasi yang sebenarnya dan prediksi model (Robeson & Willmott, 2023).

$$MSE = \frac{1}{n} \sum_i^n (\tilde{y}_i - y_i)^2 \quad (13)$$

Kemudian, *Root Mean Square Error* adalah teknik yang mudah diimplementasikan dan telah sering digunakan dalam berbagai studi yang berkaitan dengan peramalan (Sari et al., 2017).

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (\tilde{y}_i - y_i)^2} \quad (14)$$

Lalu, *Mean Absolute Error* (MAE) memberikan interpretasi yang lebih jelas dan merupakan rata - rata kesalahan yang sebenarnya (Robeson & Willmott, 2023).

$$MAE = \frac{1}{n} \sum_i^n |\tilde{y}_i - y_i| \quad (15)$$

Terakhir, *Mean Absolute Percentage Error* merupakan rata-rata *diferensiasi absolut* antara nilai peramalan dan aktual yang dinyatakan sebagai persentase nilai aktual (Riyadi, 2015).

$$MAPE = 100 \times \frac{1}{n} \sum_{i=1}^n \left| \frac{\tilde{y}_i - y_i}{y_i} \right| \quad (16)$$

Dimana \tilde{y}_i merupakan hasil peramalan; y_i merupakan nilai aktual/nilai sebenarnya; dan n adalah jumlah data.

III. HASIL DAN PEMBAHASAN

3.1 Preprocessing

Hasil *preprocessing* terhadap data *missing values* sebanyak 23 nilai dengan metode interpolasi LSTM menghasilkan data yang lengkap. Data yang semula muncul sebagai NaN, sudah diubah menjadi suatu nilai hasil prediksi dari model LSTM yang digunakan.

3.2 ARCH/GARCH

Dari eksplorasi data yang dilakukan, konsentrasi PM2.5 yang digunakan dalam penelitian ini stasioner di level. Hal ini dibuktikan dengan uji formal *Augmented Dickey-Fuller* pada data konsentrasi PM2.5 di Wilayah Jakarta Pusat dan didapatkan bahwa *p-value* sebesar 0.01.

Tabel 1. Uji ADF-Test Data Konsentrasi PM2.5 Wilayah Jakarta Pusat

Uji	p-value
ADF-Test	0.01

Kemudian, proses berlanjut dengan pencarian kandidat model terbaik dengan menerapkan metode *Box-Jenkins*. Dengan memperhatikan pola ACF dan PACF, terlihat bahwa kandidat model terbaik yaitu ARIMA(1,1,1), ARIMA(2,1,1), dan ARIMA(2,1,0). Pola ACF dan PACF mengindikasikan bahwa model harus di diferensiasi karena menunjukkan data yang tidak stasioner. Hal ini bertentangan dengan uji *Augmented Dickey-Fuller* yang signifikan. Kasus seperti ini disebabkan karena adanya heteroskedastisitas dalam data. Oleh karena itu, dilakukan percobaan dalam pembuatan model dengan memperhitungkan AIC sebagai indikator model terbaik. Kandidat model yang akan dilakukan percobaan adalah ARIMA(1,1,1), ARIMA(2,1,1), ARIMA(2,1,0), ARMA(1,1), ARMA(1,2), ARMA(2,1), dan ARMA(2,2). Dari hasil AIC diperoleh bahwa nilai AIC model ARMA(2,1) merupakan yang terkecil

dibandingkan dengan kandidat model yang lain.

Tabel 2. Tabel perbandingan nilai AIC kandidat model terbaik

Model	AIC
ARIMA(1,1,1)	1851.90
ARIMA(2,1,1)	1853.84
ARIMA(2,1,0)	1851.92
ARMA(1,1)	2081.55
ARMA(1,2)	2004.92
ARMA(2,1)	1850.51
ARMA(2,2)	1852.64

Estimasi parameter model terbaik dilakukan dengan membandingkan tingkat signifikansi yang digunakan (5%) dengan *p-value* dalam uji hipotesis. Hasil yang diperoleh bahwa parameter model ARMA(2,1) seluruhnya adalah signifikan.

Tabel 3. Tabel Uji Signifikansi Parameter model ARMA(2,1)

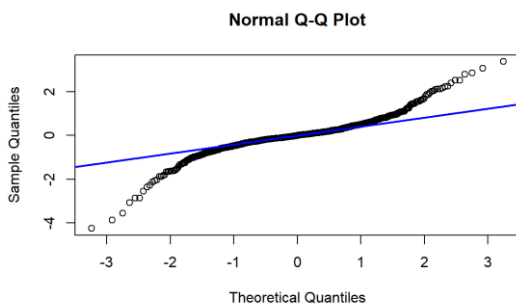
	Estimate	p-value
AR(1)	1.6963	< 2e-16
AR(2)	-0.7057	< 2e-16
MA(1)	0.1149	0.01104
Intercept	39.9723	< 2e-16

Didapatkan model rata-rata terbaik untuk meramalkan konsentrasi PM2.5 di Wilayah Jakarta Pusat adalah ARMA(2,1) dengan persamaan sebagai berikut:

$$Y_t = 39.9723 + 1.6963Y_{t-1} - 0.7057Y_{t-2} + e_t - 0.1149e_{t-1} \quad (17)$$

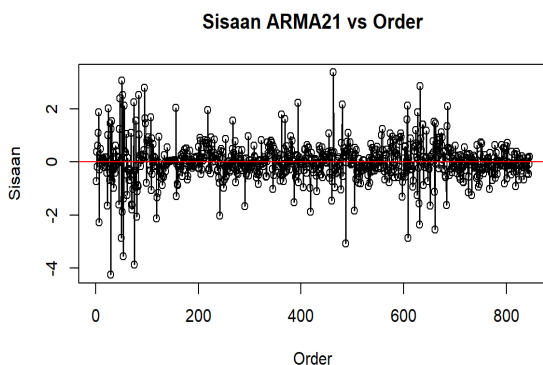
Dari model yang didapatkan, dilanjutkan pengujian residual untuk menganalisis pola residual kuadrat dengan uji korelogram dan uji

formal ARCH-LM untuk melihat heteroskedastisitas dalam data training. Pertama, dilakukan eksplorasi dengan menggunakan plot. Dari Q-Q Plot terlihat bahwa amatan cenderung menjauhi garis Q-Q plot distribusi normal, sehingga secara eksploratif dapat disimpulkan bahwa sisaan belum cukup menyebar normal.



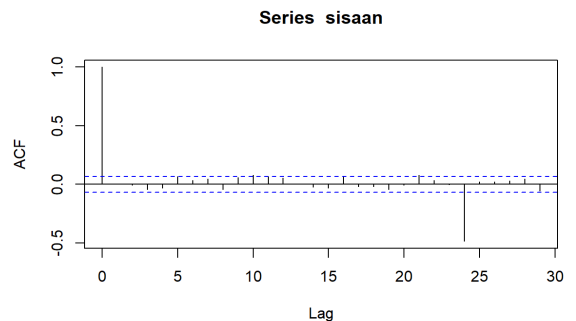
Gambar 5. Q-Q Plot residual model ARMA(2,1)

Plot residual vs order menunjukkan bahwa titik pada plot kebebasan sisaan mayoritas bergerak di sekitar titik nol. Namun, terdapat beberapa titik amatan yang terletak cukup jauh dari titik nol. Sehingga, belum dapat disimpulkan apakah terdapat autokorelasi atau tidak.

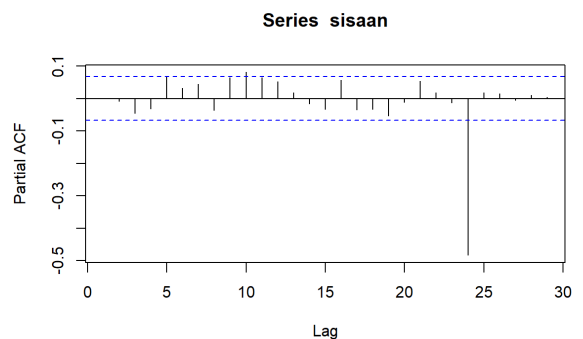


Gambar 6. Plot residual vs order model ARMA(2,1)

Berdasarkan hasil eksplorasi, baik dari plot ACF maupun plot PACF, pada keduanya terdapat garis vertikal di lag tertentu yang melebihi tinggi garis biru horizontal. Artinya, menurut kedua plot ini, terdapat autokorelasi pada model.



Gambar 7. Plot ACF model ARMA(2,1)



Gambar 8. Plot PACF model ARMA(2,1)

Untuk memperkuat dugaan, uji formal digunakan untuk mengetahui sisaan menyebar normal, sisaan saling bebas (tidak ada autokorelasi), nilai tengah sisaan sama dengan nol, dan uji formal ARCH-LM test. Hasil uji formal, didapatkan bahwa residual menyebar tidak normal, tidak terdapat autokorelasi, nilai tengah residual nol, dan terdapat heteroskedastisitas yang diuji sampai lag 15.

Tabel 4. Tabel uji formal asumsi residual

Asumsi	p-value
normalitas residual	< 2.2e-16
autokorelasi	0.9882
nilai tengah residual	0.8174

Tabel 5. Tabel uji heteroskedastisitas

Lag	p-value
1	2.293751e-16
2	5.021962e-17
3	8.312175e-18
4	2.25383e-18
5	8.048692e-18
6	2.989206e-17
7	2.380252e-16
8	2.451295e-16
9	3.722665e-16
10	1.005193e-15

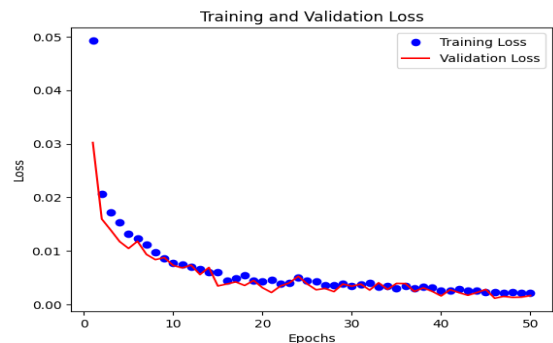
Berdasarkan hasil uji heteroskedastisitas diperoleh nilai p -value sampai lag ke-10 lebih kecil dari $\alpha = 5\%$. Dengan kata lain, dapat disimpulkan bahwa terdapat unsur heteroskedastisitas pada sisaan model ARMA(2,1). Kondisi ini dapat dikatakan bahwa terdapat indikasi efek ARCH/GARCH. Selanjutnya, dilakukan pencarian model terbaik dengan menggabungkan model ARMA(2,1) dengan kandidat model terbaik ARCH/GARCH. Model dicocokkan dengan data *training*. Dari hasil tersebut, diperoleh bahwa model terbaik, yaitu model ARMA(2,1)-GARCH(1,2) dengan nilai *Akaike Information Criterion* (AIC) sebesar

Tabel 6. Tabel kandidat model ARMA + GARCH terbaik

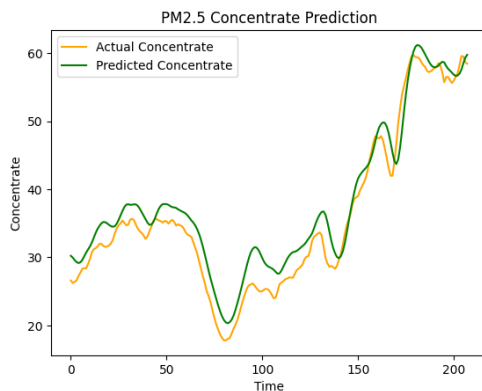
Model	AIC
ARMA(2,1) - GARCH(1,1)	1.9374
ARMA(2,1) - GARCH(1,2)	1.9109
ARMA(2,1) - GARCH(1,3)	1.9139
ARMA(2,1) - GARCH(2,1)	1.9389
ARMA(2,1) - GARCH(2,2)	1.9133
ARMA(2,1) - GARCH(2,3)	1.9163

3.3 Long Short-Term Memory (LSTM)

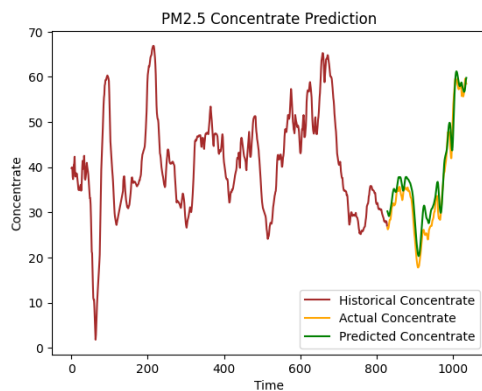
Pemodelan LSTM dilakukan dengan melihat plot data terlebih dahulu. Selanjutnya, data dibagi menjadi 80% data *training* dan 20% data *testing*. Data *training* digunakan untuk melatih model LSTM untuk belajar memahami data yang akan digunakan untuk *forecasting*. Sementara itu, data testing digunakan untuk melihat kecocokan hasil *forecasting* data *training*. Model LSTM dengan 3 lapisan di buat dan dilatih sebanyak 50 kali iterasi. Dari hasil *training*, didapatkan bahwa *training loss* dan *validation loss* semakin menurun ketika iterasi yang dilakukan semakin banyak. Hal ini mengindikasikan bahwa model yang dibangun belajar dengan baik dari data *training*.

**Gambar 9.** Training Loss dan Validation Loss model LSTM

Model kemudian digunakan untuk proses peramalan data training sebanyak 211 langkah ke depan. Hasil dari peramalan data *training* kemudian dicocokkan dengan data *testing*. Dari hasil *forecasting* yang diperoleh, terlihat bahwa hasil *forecasting* menggunakan model yang sudah dibangun menunjukkan kemiripan dari segi pola walaupun tidak sepenuhnya fit terhadap nilai sebenarnya.



Gambar 10. Plot data hasil peramalan dengan data *testing*



Gambar 11. Plot dataset dengan hasil peramalan model LSTM

3.4 Evaluasi Model

Model ARMA(2,1)-GARCH(1,2) kemudian digunakan untuk melakukan *forecasting* data training sebanyak 211 langkah ke depan lalu kemudian dievaluasi dan didapatkan hasil

evaluasi berdasarkan nilai statistik MSE, RMSE, MAE, dan MAPE sebagai berikut:

Tabel 7. Uji Performa Model ARMA(2,1)-GARCH(1,2)

Ukuran	Nilai
MSE	120.458
RMSE	10.97
MAE	7.46
MAPE	19.96

Di sisi lain, model LSTM yang dibuat juga di evaluasi menggunakan uji statistik MSE, RMSE, MAE, dan MAPE untuk melihat performa model dalam melakukan peramalan data. Nilai MSE, RMSE, MAE, dan MAPE dari model LSTM yang dibuat didapatkan dengan nilai sebagai berikut:

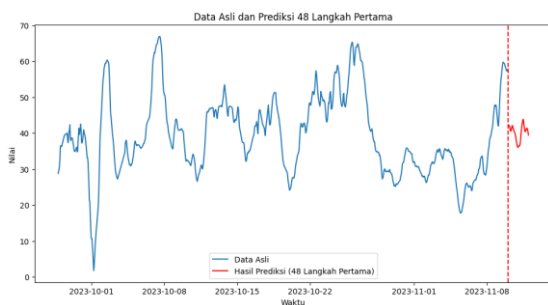
Tabel 8. Uji Performa Model LSTM

Ukuran	Nilai
MSE	7.91
RMSE	2.46
MAE	2.81
MAPE	8.07

Dari model ARMA(2,1)-GARCH(1,2) dan model LSTM, didapatkan bahwa model LSTM lebih baik untuk melakukan peramalan data konsentrasi PM2.5 di Wilayah Jakarta Pusat. Hal ini dapat dilihat dari ukuran MSE, RMSE, MAE, dan MAPE model LSTM yang dibangun lebih kecil dibandingkan dengan model ARMA(2,1)-GARCH(1,2). Selanjutnya, peramalan akan menggunakan model LSTM yang dibangun untuk meramalkan konsentrasi PM2.5 di Wilayah Jakarta Pusat beberapa jam kedepan.

3.5 Peramalan Menggunakan Algoritma Long Short-Term Memory (LSTM)

Model LSTM digunakan untuk peramalan konsentrasi PM2.5 di Wilayah Jakarta Pusat. Pada penelitian ini, model LSTM digunakan untuk meramalkan sejauh 24 jam kedepan dari data terakhir dataset, yaitu 11 November 2023 Pukul 01.00 WIB sehingga peramalan akan dilakukan hingga 12 November 2023 Pukul 01.00 WIB. Hasil dari peramalan menunjukkan adanya penurunan konsentrasi PM2.5 di Wilayah Jakarta Pusat. Namun, menurut klasifikasi dari IQAir, hasil peramalan yang dilakukan dengan model LSTM ini menunjukkan bahwa konsentrasi PM2.5 di Wilayah Jakarta Pusat berada di rentang tidak sehat bagi kelompok sensitif dan di beberapa jam berikutnya masuk ke kategori tidak sehat. Meskipun demikian, konsentrasi PM2.5 menurun jika dibandingkan dengan beberapa jam sebelumnya. Hal ini dapat terjadi karena kurangnya aktivitas kendaraan di Wilayah Jakarta Pusat pada waktu peramalan atau faktor lain yang menyebabkan konsentrasi PM2.5 berkurang seperti faktor cuaca.



Gambar 12. Hasil *forecasting* dari model LSTM yang dibangun

IV. KESIMPULAN

Salah satu polutan yang berbahaya, yaitu PM2.5 dapat menyebabkan berbagai gangguan kesehatan. Dilakukannya forecasting terhadap konsentrasi PM2.5 merupakan hal yang penting untuk

mengetahui bagaimana kondisi konsentrasi PM2.5. Dari forecasting tersebut bisa dijadikan bahan pertimbangan kebijakan dan aturan dari pemerintah terkait pengendalian polusi udara agar masyarakat memperoleh kesehatan yang lebih baik dan tercapainya visi jangka panjang tahun 2030, yaitu kota yang aman, nyaman, produktif, dan berkelanjutan.

Pada penelitian ini dilakukan peramalan konsentrasi PM2.5 menggunakan metode ARCH/GARCH dan LSTM dengan cara membandingkan performa kedua metode tersebut dalam melakukan peramalan konsentrasi PM2.5 di Wilayah Jakarta Pusat. Dalam penelitian ini nilai *Akaike Information Criterion* (AIC) digunakan sebagai indikator model terbaik. Semakin kecil nilai AIC, maka model makin baik. Dari eksplorasi data dan hasil percobaan, ARMA(2,1) merupakan model rata-rata terbaik untuk dataset yang digunakan. Di sisi lain, dataset memiliki variasi yang tidak konstan, maka dilakukan pengajuan model GARCH tambahan sebagai model varians. Dari hasil percobaan, diperoleh bahwa penggabungan model ARMA(2,1)-GARCH(1,2) memiliki nilai AIC terkecil sehingga efektif untuk melakukan prediksi PM2.5 di Wilayah Jakarta Pusat. Model tersebut digunakan untuk peramalan data testing dan di evaluasi dengan nilai MSE, RMSE, MAE, dan MAPE. Sementara itu, metode LSTM menggunakan 2 lapisan bidirectional dan 1 LSTM standar. Dari model LSTM yang dibangun, kemudian dilanjutkan peramalan data testing dan di evaluasi dengan nilai MSE, RMSE, MAE, dan MAPE.

Dengan dilakukan evaluasi terhadap *Mean Squared Error* (MSE), *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) disimpulkan bahwa *Long Short Term Memory* (LSTM) lebih baik dibanding dengan metode

GARCH didapat dari nilai *error* tersebut yang lebih kecil. Kemudian model LSTM digunakan untuk meramalkan konsentrasi PM2.5 di Wilayah Jakarta Pusat sejauh 48 langkah ke depan. Mengingat *dataset* yang digunakan adalah per-jam, maka model LSTM akan dicoba untuk meramalkan selama 2 hari kedepan. Hasil yang didapatkan yaitu konsentrasi PM2.5 di Wilayah Jakarta Pusat berada di rentang tidak sehat bagi kelompok sensitif dan di beberapa jam berikutnya masuk ke kategori tidak sehat.

DAFTAR PUSTAKA

- American Lung Association. (2023, November 2). *Who is Most Affected by Outdoor Air Pollution?*
- Badan Pusat Statistik. (2019). *Statistik Komuter Jabodetabek 2019*.
- Bové, H., Bongaerts, E., Slenders, E., Bijmens, E. M., Saenen, N. D., Gyselaers, W., Van Eyken, P., Plusquin, M., Roeflaers, M. B. J., Ameloot, M., & Nawrot, T. S. (2019). Ambient black carbon particles reach the fetal side of human placenta. *Nature Communications*, 10(1), 1–7. <https://doi.org/10.1038/s41467-019-11654-3>
- Br Perangin-angin, E. E., Bahtiar, H., Hasna, N. F., & Kartiasih, F. (2024). A VECM Approach to Assessing The Impact of Economic Growth, Livestock Production Index, and Crop Production Index on Methane Gas Emissions in Indonesia. *Jurnal Pertanian Agros*, 26(1), 4711–4732. <http://dx.doi.org/10.37159/j.p.agros.v26i1.3841>
- Cantini, F., Niccoli, L., Matarrese, D., Nicastrì, E., Stobbione, P., & Goletti, D. (2020). Baricitinib therapy in COVID-19: A pilot study on safety and clinical impact. In *Journal of Infection* (Vol. 81, Issue 2, pp. 318–356). W.B. Saunders Ltd. <https://doi.org/10.1016/j.jinf.2020.04.017>
- Chaney, R. A., Sloan, C. D., Cooper, V. C., Robinson, D. R., Hendrickson, N. R., McCord, T. A., & Johnston, J. D. (2017). Personal exposure to fine particulate air pollution while commuting: An examination of six transport modes on an urban arterial roadway. *PLoS ONE*, 12(11), 1–15. <https://doi.org/10.1371/journal.pone.0188053>
- Chen, T., Chen, F., Wang, K., Ma, X., Wei, X., Wang, W., Huang, P., Yang, D., Xia, Z., & Zhao, Z. (2021). Acute respiratory response to individual particle exposure (PM1.0, PM2.5 and PM10) in the elderly with and without chronic respiratory diseases. *Environmental Pollution*, 271, 116329. <https://doi.org/10.1016/j.envpol.2020.116329>
- Cheung, K., Daher, N., Kam, W., Shafer, M. M., Ning, Z., Schauer, J. J., & Sioutas, C. (2011). Spatial and temporal variation of chemical composition and mass closure of ambient coarse particulate matter (PM10-2.5) in the Los Angeles area. *Atmospheric Environment*, 45(16), 2651–2662. <https://doi.org/10.1016/j.atmosenv.2011.02.066>
- Com, I. (2018). *Air Pollution And Child Health*. <http://apps.who.int/bookorders>.
- Costa, L. G., Cole, T. B., Dao, K., Chang, Y. C., Coburn, J., & Garrick, J. M. (2020). Effects of air pollution on the nervous system and its possible role in neurodevelopmental and neurodegenerative disorders. *Pharmacology and Therapeutics*, 210, 107523. <https://doi.org/10.1016/j.pharmthera.2020.107523>

- Elheddad, M., Benjasak, C., Deljavan, R., Alharthi, M., & Almabrok, J. M. (2021). The effect of the Fourth Industrial Revolution on the environment: The relationship between electronic finance and pollution in OECD countries. *Technological Forecasting and Social Change*, 163(November), 120485. <https://doi.org/10.1016/j.techfore.2020.120485>
- Galbraith, J. W., & Zinde-walsh, V. (2001). Autoregression-Based Estimators for ARFIMA Models. *Cirano*.
- Grzywa-Celińska, A., Krusiński, A., & Milanowski, J. (2020). 'Smogging kills' – Effects of air pollution on human respiratory system. *Annals of Agricultural and Environmental Medicine*, 27(1), 1–5. <https://doi.org/10.26444/aaem/110477>
- Hamra, G. B., Guha, N., Cohen, A., Laden, F., Raaschou-Nielsen, O., Samet, J. M., Vineis, P., Forastiere, F., Saldiva, P., Yorifuji, T., & Loomis, D. (2014). Outdoor particulate matter exposure and lung cancer: A systematic review and meta-analysis. *Environmental Health Perspectives*, 122(9), 906–911. <https://doi.org/10.1289/ehp.1408092>
- Jayadri, B. L., Pangastuti, M., Farhan, M., & Kartiasih, F. (2024). Determinants of PM2.5 Concentration in DKI Jakarta Province: A VAR Model Approach. *Inferensi*, 7(1), 27–40. <https://doi.org/10.12962/j27213862.v7i1.19843>
- Kan, H., Chen, R., & Tong, S. (2012). Ambient air pollution, climate change, and population health in China. *Environment International*, 42(1), 10–19. <https://doi.org/10.1016/j.envint.2011.03.003>
- Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, 64(5), 402–406. <https://doi.org/10.4097/kjae.2013.64.5.402>
- Kartiasih, F., & Setiawan, A. (2020). Aplikasi Error Correction Mechanism dalam Analisis Dampak Pertumbuhan Ekonomi, Konsumsi Energi dan Perdagangan Internasional Terhadap Emisi CO2 di Indonesia. *Media Statistika*, 13(1), 104–115. <https://doi.org/10.14710/medstat.13.1.104-115>
- Kitamura, H., Dahlan, A. V., Tian, Y., Shimaoka, T., Yamamoto, T., & Takahashi, F. (2018). Impact of secondary generated minerals on toxic element immobilization for air pollution control fly ash of a municipal solid waste incinerator. *Environmental Science and Pollution Research*, 25(21), 20700–20712. <https://doi.org/10.1007/s11356-018-1959-5>
- Leachi, H. F. L., Marziale, M. H. P., Martins, J. T., Aroni, P., Galdino, M. J. Q., & Ribeiro, R. P. (2020). Polycyclic aromatic hydrocarbons and development of respiratory and cardiovascular diseases in workers. *Revista Brasileira de Enfermagem*, 73(3), 1–8. <https://doi.org/10.1590/0034-7167-2018-0965>
- Lei, R., Zhu, F., Cheng, H., Liu, J., Shen, C., Zhang, C., Xu, Y., Xiao, C., Li, X., Zhang, J., Ding, R., & Cao, J. (2019). Short-term effect of PM2.5/O3 on non-accidental and respiratory deaths in highly polluted area of China. *Atmospheric Pollution Research*, 10(5), 1412–1419.

- <https://doi.org/10.1016/j.apr.2019.03.013>
- Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A Review. *Frontiers in Public Health*, 8(February), 1–13. <https://doi.org/10.3389/fpubh.2020.00014>
- Manousakas, M., Papaefthymiou, H., Diapouli, E., Migliori, A., Karydas, A. G., Bogdanovic-Radovic, I., & Eleftheriadis, K. (2017). Assessment of PM2.5 sources and their corresponding level of uncertainty in a coastal urban area using EPA PMF 5.0 enhanced diagnostics. *Science of the Total Environment*, 574, 155–164. <https://doi.org/10.1016/j.scitotenv.2016.09.047>
- Martelletti, L., & Martelletti, P. (2020). Air Pollution and the Novel Covid-19 Disease: a Putative Disease Risk Factor. *SN Comprehensive Clinical Medicine*, 2(4), 383–387. <https://doi.org/10.1007/s42399-020-00274-4>
- Masum, S., Liu, Y., & Chiverton, J. (2018). Multi-step time series forecasting of electric load using machine learning models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10841 LNAI*. Springer International Publishing. https://doi.org/10.1007/978-3-319-91253-0_15
- Mukherjee, A., & Agrawal, M. (2017). World air particulate matter: sources, distribution and health effects. *Environmental Chemistry Letters*, 15(2), 283–309. <https://doi.org/10.1007/s10311-017-0611-9>
- Nor, M. E., Safuan, H. M., Shab, N. F. M., Asrul, M., Abdullah, A., Mohamad, N. A. I., & Lee, M. H. (2017). Neural network versus classical time series forecasting models. *AIP Conference Proceedings*, 1842. <https://doi.org/10.1063/1.4982865>
- Pini, L., Giordani, J., Gardini, G., Concoreggi, C., Pini, A., Perger, E., Vizzardi, E., Di Bona, D., Cappelli, C., Ciarfaglia, M., & Tantucci, C. (2021). Emergency department admission and hospitalization for COPD exacerbation and particulate matter short-term exposure in Brescia, a highly polluted town in northern Italy. *Respiratory Medicine*, 179(February), 106334. <https://doi.org/10.1016/j.rmed.2021.106334>
- Pribadi, W., & Kartiasih, F. (2020). Environmental Quality and Poverty Assessment in Indonesia. *Jurnal Pengelolaan Sumberdaya Alam Dan Lingkungan (Journal of Natural Resources and Environmental Management)*, 10(1), 89–97. <https://doi.org/10.29244/jpsl.10.1.89-97>
- Puri, P., Nandar, S. K., Kathuria, S., & Ramesh, V. (2017). Effects of air pollution on the skin: A review. *Indian Journal of Dermatology, Venereology and Leprology*, 83(4), 415–423. <https://doi.org/10.4103/0378-6323.199579>
- Riyadi, S. (2015). Aplikasi Peramalan Penjualanobat Menggunakan Metodepemulusan (Studi Kasus: Instalasi Farmasi Rsud Dr Murjani). *Seminar Nasional Teknologi Informasi Dan Multimedia 2015*, 1, 1–6.
- Robeson, S. M., & Willmott, C. J. (2023). Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. *PLoS ONE*,

- 18(2 February), 1–8.
<https://doi.org/10.1371/journal.pone.0279774>
- Sahoo, B. B., Jha, R., Singh, A., & Kumar, D. (2019). Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica*, 67(5), 1471–1481.
<https://doi.org/10.1007/s11600-019-00330-1>
- Sari, N. R., Mahmudy, W. F., Wibawa, A. P., & Sonalitha, E. (2017). Enabling external factors for inflation rate forecasting using fuzzy neural system. *International Journal of Electrical and Computer Engineering*, 7(5), 2746–2756.
<https://doi.org/10.11591/ijece.v7i5.pp2746-2756>
- Shafiei, S., & Salim, R. A. (2014). Non-renewable and renewable energy consumption and CO2 emissions in OECD countries: A comparative analysis. *Energy Policy*, 66, 547–556.
<https://doi.org/10.1016/j.enpol.2013.10.064>
- Shahrbafe, M. A., Akbarzadeh, M. A., Tabary, M., & Khaheshi, I. (2021). Air Pollution and Cardiac Arrhythmias: A Comprehensive Review. *Current Problems in Cardiology*, 46(3).
<https://doi.org/10.1016/j.cpcardiol.2020.100649>
- Shi, J., Jain, M., & Narasimhan, G. (2022). *Time Series Forecasting (TSF) Using Various Deep Learning Models*. June.
- Švédová, B., Raclavská, H., Kucbel, M., Růžicková, J., Raclavský, K., Koliba, M., & Juchelková, D. (2020). Concentration variability of water-soluble ions during the acceptable and exceeded pollution in an industrial region. *International Journal of Environmental Research and Public Health*, 17(10).
<https://doi.org/10.3390/ijerph17103447>
- Swardanasuta, I. B. P., Sandy, N. R. K., Rohmah, N. A., Arindah, Y., & Kartiasih, F. (2024). The Effect of Industrial Value Added, Energy Consumption, Food Crop Production, and Air Temperature on Greenhouse Gas Emissions in Indonesia: A Time Series Analysis Approach. *Jurnal Pertanian Agros*, 26(1), 4848–4865.
<http://dx.doi.org/10.37159/j.agros.v26i1.3876>
- Thangavel, P., Park, D., & Lee, Y. C. (2022). Recent Insights into Particulate Matter (PM2.5)-Mediated Toxicity in Humans: An Overview. *International Journal of Environmental Research and Public Health*, 19(12).
<https://doi.org/10.3390/ijerph19127511>
- Wang, Y., Li, C., Zhang, X., Kang, X., Li, Y., Zhang, W., Chen, Y., Liu, Y., Wang, W., Ge, M., & Du, L. (2021). Exposure to PM2.5 aggravates Parkinson's disease via inhibition of autophagy and mitophagy pathway. *Toxicology*, 456(January), 152770.
<https://doi.org/10.1016/j.tox.2021.152770>
- WHO. (2023). *World Health Organization: Air Pollution*.
- Wilson, G. T. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. *Journal of Time Series Analysis*, 37(5), 709–711. <https://doi.org/10.1111/jtsa.12194>
- Wilson, W. E., & Suh, H. H. (1997). Fine particles and coarse particles: Concentration relationships relevant to

- epidemiologic studies. *Journal of the Air and Waste Management Association*, 47(12), 1238–1249. <https://doi.org/10.1080/10473289.1997.10464074>
- Wyatt, L. H., Weaver, A. M., Moyer, J., Schwartz, J. D., Di, Q., D.-S., D., Cascio, W. E., & Ward-Caviness, C. K. (2022). Short-term PM2.5 exposure and early-readmission risk: A retrospective cohort study in North Carolina heart failure patients. *Am. Heart J*, 248, 130–138.
- Xie, G., Sun, L., Yang, W., Wang, R., Shang, L., Yang, L., Qi, C., Xin, J., Yue, J., & Chung, M. C. (2021). Maternal exposure to PM2.5 was linked to elevated risk of stillbirth. *Chemosphere*, 283, 131169. <https://doi.org/10.1016/j.chemosphere.2021.131169>
- Yen, N. Y., Chang, J. W., Liao, J. Y., & Yong, Y. M. (2020). Analysis of interpolation algorithms for the missing values in IoT time series: a case of air quality in Taiwan. *Journal of Supercomputing*, 76(8), 6475–6500. <https://doi.org/10.1007/s11227-019-02991-7>
- Yuan, H., Xu, G., Yao, Z., Jia, J., & Zhang, Y. (2018). Imputation of missing data in time series for air pollutants using long short-term memory recurrent neural networks. *UbiComp/ISWC 2018 - Adjunct Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2018 ACM International Symposium on Wearable Computers*, 1293–1300. <https://doi.org/10.1145/3267305.3274648>
- Zhang, G., Rui, X., & Fan, Y. (2018). Critical review of methods to estimate PM2.5 concentrations within specified research region. *Canadian Historical Review*, 7(9). <https://doi.org/10.3390/ijgi7090368>
- Zhang, L., Yang, Y., Li, Y., Qian, Z. (Min), Xiao, W., Wang, X., Rolling, C. A., Liu, E., Xiao, J., Zeng, W., Liu, T., Li, X., Yao, Z., Wang, H., Ma, W., & Lin, H. (2019). Short-term and long-term effects of PM2.5 on acute nasopharyngitis in 10 communities of Guangdong, China. *Science of the Total Environment*, 688, 136–142. <https://doi.org/10.1016/j.scitotenv.2019.05.470>
- Zhang, Y., Ding, Z., Xiang, Q., Wang, W., Huang, L., & Mao, F. (2020). Short-term effects of ambient PM1 and PM2.5 air pollution on hospital admission for respiratory diseases: Case-crossover evidence from Shenzhen, China. *International Journal of Hygiene and Environmental Health*, 224(September), 113418. <https://doi.org/10.1016/j.ijheh.2019.11.001>
- Zhao, C. N., Xu, Z., Wu, G. C., Mao, Y. M., Liu, L. N., Qian-Wu, Dan, Y. L., Tao, S. S., Zhang, Q., Sam, N. B., Fan, Y. G., Zou, Y. F., Ye, D. Q., & Pan, H. F. (2019). Emerging role of air pollution in autoimmune diseases. *Autoimmunity Reviews*, 18(6), 607–614. <https://doi.org/10.1016/j.autrev.2018.12.010>