

**PENILAIAN KOMPARATIF MODEL SARIMA DAN LSTM UNTUK
KNOWLEDGE DISCOVERY AIR QUALITY INDEX PADA KOTA
GURUGRAM**

SKRIPSI

Program Studi Sistem Informasi

Jenjang Sarjana



Oleh :

Nanda Aulia Sofiah

NIM 09031182126017

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN

SKRIPSI

**PENILAIAN KOMPARATIF MODEL SARIMA DAN LSTM UNTUK
KNOWLEDGE DISCOVERY AIR QUALITY INDEX PADA KOTA
GURUGRAM**

Sebagai salah satu syarat untuk penyelesaian
studi di program studi Sistem Informasi S1

Oleh :

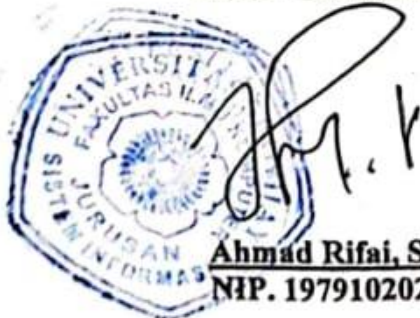
Nanda Aulia Sofiah 09031182126017

Mengetahui,

Palembang, 12 November 2024

Ketua Jurusan Sistem Informasi

Pembimbing



Ahmad Rifai, S.T., M.T.
NIP. 19791020201021003



Ken Ditha Tania, M.Kom., PhD.
NIP. 198507182012122003

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini :

Nama : Nanda Aulia Sofiah

NIM : 09031182126017

Program Studi : Sistem Informasi Reguler

Judul Skripsi : Penilaian Komparatif Model SARIMA dan LSTM untuk
Knowledge Discovery Air Quality Index pada Kota
Gurugram

Hasil Pengecekan iThenticate/Turnitin : 1%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya saya sendiri dan bukan penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan skripsi ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 12 November 2024



Nanda Aulia Sofiah
NIM. 09031182126017

HALAMAN PERSETUJUAN

Telah accepted jurnal di *International Conference on Electrical Engineering and Computer Science 2024* pada :

Hari : Sabtu

Tanggal : 03 Agustus 2024

Nama : Nanda Aulia Sofiah

NIM : 09031182126017

Judul : *A Comparative Assessment SARIMA and LSTM Models for the Gurugram Air Quality Index's Knowledge Discovery*

Tim Pembimbing :

2. Pembimbing : Ken Ditha Tania, M.Kom., PhD.



Mengetahui,

Ketua Jurusan Sistem Infromasi



Ahmad Rifai, S.T., M.T.
NIP. 19791020201021003

HALAMAN PERSEMBAHAN DAN MOTTO

"Indeed, with hardship [will be] ease."

(Quran 94:6)

"Believe you can and you're halfway there."

Theodore Roosevelt

Skripsi ini dipersembahkan kepada :

- ❖ Allah Subhanahu Wata'ala, sebagai bentuk pemenuhan amanah kedua orang tua dan upaya menuntut ilmu dunia untuk kebutuhan akhirat kelak.
- ❖ Keluarga penulis yang tercinta.
- ❖ Dosen Pembimbing penulis, Ibu Ken Ditha Tania, M.Kom., PhD.
- ❖ Sahabat dan teman-teman seperjuangan.
- ❖ Almamaterku, Universitas Sriwijaya.

KATA PENGANTAR



Atas bantuan dan kasih sayang Allah subhanahu wata'ala, penulis berhasil menyelesaikan skripsi dengan judul " **Penilaian Komparatif Model SARIMA dan LSTM untuk Knowledge Discovery Air Quality Index pada Kota Gurugram**". Salam dan shalawat selalu diucapkan kepada junjungan besar Nabi Muhammad Shallallahu Alaihi Wassalam, yang memimpin umat Islam menuju kemajuan dan kemakmuran. Semoga kita dapat tetap mengikuti ajarannya sampai kita meninggal dunia.

Tugas akhir ini merupakan salah satu syarat untuk menyelesaikan pendidikan S1 di Jurusan Sistem Informasi di Fakultas Ilmu Komputer Universitas Sriwijaya. Penulis menyadari bahwa terdapat banyak tantangan yang dihadapi selama proses penyusunan tugas akhir ini yang disebabkan oleh berbagai faktor internal dan eksternal, kelemahan, serta masalah yang penulis alami sendiri. Namun, berkat bimbingan, bantuan, doa, dan semangat dari berbagai pihak, penulis dapat mengatasi dan menyelesaikan tugas akhir ini. Oleh karena itu, penulis ingin menyampaikan rasa hormat dan terima kasih kepada pihak-pihak berikut.

1. Allah SWT, Yang Maha Kuasa, atas rahmat, nikmat, dan karunia-Nya, seperti kesehatan, kesempatan, dan kemudahan, sehingga penulis dapat menyelesaikan skripsi ini.
2. Orang tua penulis yang selalu mendoakan, mendukung, dan merawat penulis dengan penuh kasih sayang. Seluruh upaya yang telah dilakukan

selama ini membantu penulis menjalani hidup yang baik. Diikuti dengan semangat dan dorongan yang diberikan orang tua penulis untuk membuat penulis bertahan hingga saat ini dan menyelesaikan studi S1 penulis. Terima kasih banyak atas pengorbanan yang tidak terhitung yang diberikan oleh orang tua penulis.

3. Bapak Prof. Dr. Erwin, S.Si., M.Si selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Ahmad Rifai, M.T. selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Ken Ditha Tania, M.Kom., PhD., selaku Dosen Pembimbing Tugas Akhir penulis, yang telah memberikan bantuan sepanjang proses penyelesaian tugas akhir ini. Penulis sangat berterima kasih atas informasi, gagasan, dan saran yang telah diberikan.
6. Mba Ayu selaku Admin Program Studi yang telah membantu proses pengurusan berkas dan memberikan informasi terkait perkuliahan.
7. Seluruh dosen dan karyawan yang telah membantu dan memberikan banyak pengetahuan dan arahan selama penulis menempuh pendidikan di Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Teman-teman seperjuangan jurusan Sistem Informasi angkatan 2021
9. Sahabat-sahabat saya di ABDH, yaitu Aida, Fanny, Hestiana, Jessica, dan Vanessa yang sudah banyak memberikan bantuan dan dukungan selama penulis menjalani perguruan tinggi. *Starting from the first years, I have never thought that we could make our own history. But overtime we uncover our new development even in our lowest. We conquered so many*

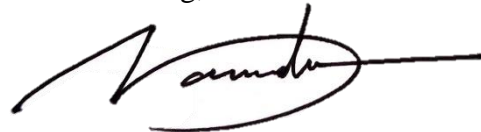
obstacles together and I can not ask for more. The support and love that you all have given really helped me get through this much better. I will forever be lucky for having you all as my friends and will certainly cherish these moments.

10. Kak Dina, kakak tingkat penulis di jurusan Sistem Infromasi, yang telah sangat membantu penulis selama masa perkuliahan. *The purest and nicest person that I've known. These words may never be enough to show you how grateful I am for having you as my senior. You definitely are the definition of loveable, genuine, and full of kindness. I am very grateful for the help, support, and other little things that you've given to me.*

11. Mechi, grup teman-teman penulis sejak masa SMA, yang terus memberikan dukungan dan apresiasi kepada penulis. *You guys are definitely one of a few things that I would never forget.*

Penulis menyadari bahwa skripsi ini masih perlu ditingkatkan karena belum sempurna. Oleh karena itu, penulis sangat terbuka untuk menerima kritik dan saran yang membantu untuk membuat skripsi lebih baik di masa depan. Penulis juga berharap skripsi ini akan membantu penelitian terkait dan menambah pengetahuan.

Palembang, 12 November 2024



Nanda Aulia Sofiah
NIM. 09031182126017

**PENILAIAN KOMPARATIF MODEL SARIMA DAN LSTM UNTUK
KNOWLEDGE DISCOVERY AIR QUALITY INDEX PADA KOTA
GURUGRAM**

Oleh

**Nanda Aulia Sofiah
09031182126017**

ABSTRAK

Tingkat polusi udara di Gurugram, sebuah pusat kota India, meningkat sebagai akibat dari urbanisasi yang cepat dan aktivitas industri yang meningkat. Kondisi kesehatan jangka panjang seperti kanker paru-paru dan penyakit kardiovaskular dapat disebabkan oleh polusi udara, yang bertanggung jawab atas 25% dari semua kematian di negara-negara terbelakang. Untuk mengimplementasikan langkah-langkah kesehatan masyarakat dan lingkungan yang efektif, estimasi yang akurat dari *Air Quality Index* (AQI) sangat penting. Tujuan ini mengevaluasi akurasi prediksi *Air Quality Index* (AQI) di Kota Gurugram dengan menggunakan dua algoritma pembelajaran mesin, yaitu SARIMA dan LSTM. Studi ini bertujuan untuk melakukan knowledge discovery dengan membandingkan akurasi model SARIMA dan LSTM dalam memprediksi *Air Quality Index* (AQI) Kota Gurugram yang dapat digunakan untuk menjaga kesehatan masyarakat dan mendeteksi keadaan darurat pada tahap awal. Pengetahuan yang ditemukan menunjukkan bahwa model LSTM memiliki tingkat variabilitas kesalahan yang lebih tinggi dengan *Root Mean Square Error* (RMSE) sebesar 63,163 dibandingkan dengan model SARIMA yang memiliki *Root Mean Square Error* (RMSE) senilai 61.999. Namun, model LSTM memiliki nilai *Mean Absolute Error* (MAE) yang lebih kecil, yaitu senilai 40,948 dibandingkan dengan model SARIMA yang memiliki nilai *Mean Absolute Error* (MAE) sebesar 45,758.

Kata Kunci: *Air Quality Index*; Gurugram; Prediksi; *Machine Learning Algorithm*; *Knowledge Discovery*.

A COMPARATIVE ASSESSMENT SARIMA AND LSTM MODELS FOR THE GURUGRAM AIR QUALITY INDEX'S KNOWLEDGE DISCOVERY

Oleh

Nanda Aulia Sofiah
09031182126017

ABSTRACT

Gurugram, an urban center in India, experiences elevated levels of air pollution as a result of rapid urbanization and intensified industrial operations. Air pollution, responsible for 25% of all deaths in underdeveloped countries, can lead to long-term health conditions such as lung cancer and cardiovascular disease. Precise forecasting of the Air Quality Index (AQI) is crucial for the implementation of effective environmental and public health measures. The objective of this project is to evaluate the precision of Gurugram's Air Quality Index (AQI) predictions by employing two machine learning algorithms, namely SARIMA and LSTM. This study aims to gain knowledge discovery by comparing the accuracy of SARIMA and LSTM models in forecasting AQI for Gurugram that can be utilized for safeguarding public health and detecting potential emergencies at an early stage. Discoveries indicate that although the LSTM model has a higher degree of error variability (with a root mean square error of 63.163 compared to 61.999 for SARIMA), it surpasses SARIMA in terms of average forecast accuracy (with a mean absolute error of 40.948 versus 45.758).

Keywords: *Air Quality Index; Gurugram; Forecasting; Machine Learning Algorithm; Knowledge Discovery.*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN DAN MOTTO	v
KATA PENGANTAR.....	vi
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Landasan Teori	7
2.1.1 <i>Air Quality Index (AQI)</i>	7
2.1.2 <i>Seasonal Autoregressive Integrated Moving Average (SARIMA)</i> ... 8	
2.1.3 <i>Long Short-Term Memory (LSTM)</i>	9
2.1.4 <i>Knowledge Discovery</i>	9
2.2 Penelitian Terdahulu.....	10
BAB III METODOLOGI PENELITIAN	13
3.1 Objek Penelitian	13
3.2 Alat dan Bahan	13
3.3 Tahapan Penelitian	14
3.4 Tahap Perolehan Data	15
3.5 Tahap Eksplorasi Data.....	16
3.6 Tahap Pembersihan Data.....	16
3.7 Tahap Pembuatan Model.....	17

3.8 Tahap Evaluasi	20
BAB IV HASIL DAN PEMBAHASAN	21
4.1 Hasil Evaluasi Model	21
4.2 Hasil Perbandingan Prediksi Model	22
BAB V PENUTUP.....	25
5.1 Kesimpulan.....	25
5.2 Saran.....	26
DAFTAR PUSTAKA	27

DAFTAR GAMBAR

Gambar 3.1 Kerangka Kerja SEMMA	14
Gambar 3.2 Importing Dataset	15
Gambar 3.3 Visualisasi Data AQI di Kota Gurugram.....	16
Gambar 3.4 Pengecekan dan Pengisian Data yang Hilang.....	16
Gambar 3.5 Hasil Seasonal Decompose Data AQI	17
Gambar 3.6 Hasil Tes ADF Data AQI	18
Gambar 3.7 Plot ACF dan PACF Data AQI.....	18
Gambar 3.8 Kode Program Model SARIMA	19
Gambar 3.9 Kode Program Model LSTM.....	19
Gambar 3.10 Kode Program Evaluasi Model.....	20
Gambar 4.1 Perbandingan Hasil Prediksi Model SARIMA.....	22
Gambar 4.2 Perbandingan Hasil Prediksi Model LSTM.....	23

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	10
Tabel 4.1 Hasil Evaluasi Model	21

DAFTAR LAMPIRAN

Lampiran 1 Surat Kesedian Membimbing	A-1
Lampiran 2 Form Pengajuan Topik Skripsi	B-1
Lampiran 3 Surat Keputusan Pembimbing Tugas Akhir	C-1
Lampiran 4 Surat Keterangan Pengecekan Similarity	D-1
Lampiran 5 Hasil Pengecekan Similarity	E-1
Lampiran 6 Kartu Konsultasi Skripsi	F-1
Lampiran 7 Logbook Dosen Pembimbing Skripsi	G-1
Lampiran 8 Form Desk Evaluasi.....	H-1
Lampiran 9 Sampel Data Penelitian	I-1
Lampiran 10 Bukti Korespondensi.....	J-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gurugram merupakan sebuah kota di India utara di sebelah barat daya New Delhi. Kota terbesar kedua di Haryana ini telah mengalami pertumbuhan yang signifikan dalam berbagai aspek pembangunan dalam beberapa tahun terakhir, terutama dalam hal kota dan sektor bisnis. Gurugram, Kota Milenium India, mengalami urbanisasi yang cepat. Investor yang kuat secara finansial dari luar negeri dan pengembang real estate yang baru berkembang menjadi bagian dari komunitas ini (Pramanik et al., 2021).

Peristiwa ini mengubah sebuah desa menjadi sebuah kota kontemporer dengan gedung pencakar langit yang menjulang tinggi, pusat perbelanjaan yang luas, distrik komersial yang padat, pusat layanan multinasional, dan jaringan transportasi yang sangat baik. Namun, peningkatan yang secara tiba-tiba ini telah menyebabkan masalah terkait kemacetan lalu lintas, genangan air, dan polusi udara (Guptha et al., 2022). Selama beberapa tahun terakhir, kota metropolitan ini telah mengalami banjir perkotaan yang sering, yang telah menyebabkan kerusakan yang signifikan.

Kerusakan tersebut di antaranya ialah banjir di jalan raya, termasuk Jalan Raya Nasional, dan dampak buruk pada lalu lintas jalan raya seperti polusi suara, kemacetan, kecelakaan di jalan raya, polusi udara, serta dampak negatif pada kesehatan dan lingkungan akibat kualitas udara yang buruk (Saxena dan Choudhury, 2022). Organisasi Kesehatan Dunia (WHO) menempatkan Gurugram

sebagai kota paling tercemar di dunia pada tahun 2019, dengan emisi mobil dan bangunan sebagai penyebab utama polusi (Shubhender dan Berwal, 2024).

Menurut WHO, India merupakan negara yang berkembang pesat dengan populasi yang terus meningkat dan menderita polusi udara yang parah. Di antara 20 kota dengan tingkat polusi udara terburuk di dunia, 13 di antaranya berada di India (Kaur dan Pandey, 2021). Meningkatnya polusi udara di sebagian besar kota-kota besar di India selama beberapa dekade terakhir berdampak terhadap kesehatan manusia. Meningkatnya kegiatan industri dan pembangunan berkontribusi terhadap polusi udara yang parah, yang menyumbang sekitar 25% dari semua kematian di negara-negara berkembang seperti yang dilaporkan oleh WHO (Indraj and Warpa 2024).

Pada tahun 2019, polusi udara di India mengakibatkan 1,67 juta kematian, menyumbang 17,8% dari total kematian. Polusi materi partikulat adalah penyebab utamanya (0,98 juta), diikuti oleh polusi udara dalam rumah (0,61 juta), dengan interval ketidakpastian 95% antara 1,42 dan 1,92 juta (Pandey et al., 2021). India bertanggung jawab atas sekitar 25% dari tiga juta kematian global yang disebabkan oleh paparan PM_{2.5}, atau materi partikulat kecil dengan diameter kurang dari 2,5 μm (Bhar et al. 2024).

Sebuah studi menunjukkan bahwa paparan polusi udara dalam jangka pendek dan jangka panjang berkorelasi dengan peningkatan jumlah kematian akibat penyakit kardiovaskular dan kejadian penyakit ini (Schaidhauer et al., 2024). Karena ukurannya yang kecil, polusi dapat dengan mudah memasuki bronkiolar dan alveolar paru-paru untuk menyebabkan kondisi pernapasan seperti asma, batuk,

dan sesak napas (Singh et al. 2023). Risiko kematian dini dapat meningkat jika terpapar polusi udara dalam jangka waktu yang lama (Natarajan et al. 2024). Jadi, bahkan orang yang sehat pun rentan terhadap penyakit dan kematian akibat udara kotor. Sangat penting untuk mengidentifikasi dan menangani wilayah yang sangat terkontaminasi dengan polusi udara (Janarthanan et al. 2021).

Salah satu upaya besar untuk mengurangi efek polusi udara pada manusia adalah sistem pemantauan kualitas udara. *Air Quality Index* (AQI) adalah perkiraan nilai yang dihubungkan dengan jumlah berbagai polutan udara ke dalam bentuk matematis untuk menghitung nilai kualitas udara, yang dapat digunakan untuk mendukung studi kesehatan manusia dan analisis dampak lingkungan (Ravindra et al., 2024). Tindakan ini dapat memungkinkan pemerintah untuk menanggapi perubahan pengetahuan tentang kimia atmosfer dan efeknya terhadap kesehatan masyarakat dan lingkungan (Horn dan Dasgupta, 2024).

Metode prediksi AQI akan memberikan dasar untuk pemantauan lingkungan untuk mengontrol dan mengurangi polusi udara di masa depan. Model *machine learning* untuk memprediksi polusi udara telah digunakan dalam beberapa tahun terakhir untuk menilai tingkat polusi udara di berbagai lokasi (Aram et al. 2024). Sebuah model *Seasonal Auto Regressive Integrated Moving Average* (SARIMA) dapat dengan akurat menggambarkan deret waktu dengan ketidakstasioneritasan periodik sederhana (Minh et al. 2024).

Sebuah penelitian menunjukkan bahwa model SARIMA dapat dengan akurat memperkirakan jumlah PM10 di Bangkok (Bunnag 2024). Selain itu, dalam memprediksi tingkat AQI, model SARIMA lebih baik daripada model statistik

konvensional (Manikandan et al. 2023). Hal ini menunjukkan kekuatan AI untuk prediksi dan peramalan AQI. Dalam pembelajaran mendalam, *Long Short-Term Memory* (LSTM) telah menjadi metode yang populer. Sebuah penelitian yang menggunakan model LSTM menunjukkan bahwa, dengan RMSE terendah, LSTM adalah algoritma yang paling efisien untuk mengantisipasi kualitas udara (Yan et al. 2021).

Knowledge discovery adalah istilah yang luas yang mencakup *data mining*, yang melibatkan penggunaan teknik-teknik khusus pemrosesan data untuk mengekstrak informasi yang berharga (Shu and Ye 2023). Penemuan dalam bidang manajemen kualitas udara memungkinkan pertimbangan yang lebih baik dan tindakan proaktif untuk mengurangi dampak negatif polusi udara terhadap kesehatan manusia dan lingkungan. Selain itu, pemahaman hasil prediktif menjadi lebih baik dengan teknik ini. Meskipun terdapat kemajuan signifikan dalam teknologi dan manajemen pengendalian polusi udara, masih sulit untuk mencapai hasil yang diinginkan (Gulia et al. 2022). Hal ini menunjukkan bahwa kebijakan dan teknologi saat ini harus ditingkatkan untuk mengatasi kesenjangan penelitian tentang masalah polusi udara di India.

Prediksi *Air Quality Index* (AQI) menunjukkan informasi mengenai kualitas udara dan dapat membantu pihak terkait mengambil tindakan pencegahan untuk melindungi diri dan lingkungan. Pembuat kebijakan juga dapat menggunakan prediksi ini untuk mengurangi polusi udara. Oleh karena itu, tujuan utama dari penelitian ini adalah untuk menemukan pengetahuan melalui perbandingan akurasi prediksi AQI serta menentukan model yang lebih baik untuk memprediksi AQI di Gurugram dengan menggunakan dua model prediktif, yaitu SARIMA dan LSTM.

Berdasarkan penjelasan penulis di atas, penulis ingin melakukan penelitian dengan judul “**PENILAIAN KOMPARATIF MODEL SARIMA DAN LSTM UNTUK KNOWLEDGE DISCOVERY AIR QUALITY INDEX PADA KOTA GURUGRAM**”.

1.2 Rumusan Masalah

Berdasarkan penjelasan di atas, tercipta rumusan masalah sebagai berikut :

1. Model apa yang lebih akurat untuk memprediksi AQI di Kota Gurugram di antara model SARIMA dan model LSTM?
2. Pengetahuan apa yang dapat ditemukan dengan membandingkan performa model SARIMA dan model LSTM dalam memprediksi AQI di Kota Gurugram?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Mengetahui model yang lebih akurat untuk memprediksi AQI di Kota Gurugram di antara model SARIMA dan model LSTM.
2. Melakukan *knowledge discovery* dengan membandingkan performa model SARIMA dan model LSTM dalam memprediksi AQI di Kota Gurugram.

1.4 Manfaat Penelitian

1. Penulis dapat memberikan pengetahuan model yang lebih sesuai untuk memprediksi AQI di Kota Gurugram.

2. Hasil dari penelitian dapat memberikan informasi mengenai kualitas udara di Kota Gurugram sehingga dapat digunakan untuk mengembangkan kebijakan mitigasi polusi udara.
3. Penelitian dapat menjadi referensi tentang efektivitas model prediktif SARIMA dan LSTM dalam memprediksi AQI yang dapat digunakan dalam penelitian lebih lanjut.

1.5 Batasan Masalah

1. Penelitian hanya berfokus pada dua model prediksi, yaitu SARIMA dan LSTM.
2. Data yang digunakan terbatas pada wilayah Gurugram, sehingga hasil penelitian mungkin tidak dapat digeneralisasi ke kota lain dengan kondisi yang berbeda.
3. Penelitian berfokus pada performa model dalam memprediksi AQI dan tidak membahas secara mendalam polutan spesifik yang berkontribusi pada kualitas udara.

DAFTAR PUSTAKA

- Agyare, Sampson, Benjamin Odoi, and Eric Neebo Wiah. 2024. "Predicting Petrol and Diesel Prices in Ghana, A Comparison of ARIMA and SARIMA Models." *Asian Journal of Economics, Business and Accounting* 24(5):594–608.
- Andrade, Chittaranjan. 2019. "The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives." *Indian Journal of Psychological Medicine* 41:210. doi: 10.4103/IJPSYM.IJPSYM_193_19.
- Aram, S. A., E. A. Nketiah, B. M. Saalidong, H. Wang, A. R. Afitiri, A. B. Akoto, and P. O. Lartey. 2024. "Machine Learning-Based Prediction of Air Quality Index and Air Quality Grade: A Comparative Analysis." *International Journal of Environmental Science and Technology* 21(2):1345–60.
- Atoui, Alya, Kamal Slim, Samir Abbad Andaloussi, Régis Moilleron, and Zaher Khraibani. 2022. "Time Series Analysis and Forecasting of the Air Quality Index of Atmospheric Air Pollutants in Zahleh, Lebanon." *Atmospheric and Climate Sciences* 12(4):728–49.
- Bhar, Soumyajit, Sharachandra Lele, Jihoon Min, and Narasimha D. Rao. 2024. "Water, Air Pollution and Carbon Footprints of Conspicuous/Luxury Consumption in India." *Ecological Economics* 218:108104. doi: <https://doi.org/10.1016/j.ecolecon.2024.108104>.
- Bunnag, Tanattrin. 2024. "Forecasting PM10 Caused by Bangkok's Leading Greenhouse Gas Emission Using the SARIMA and SARIMA-GARCH

Model.” *International Journal of Energy Economics and Policy* 14(1):418–26.

Cican, Grigore, Adrian-Nicolae Buturache, and Radu Mirea. 2023. “Applying Machine Learning Techniques in Air Quality Prediction—A Bucharest City Case Study.” *Sustainability* 15(11):8445.

Gulia, Sunil, Nidhi Shukla, Lavanya Padhi, Parthaa Bosu, S. K. Goyal, and Rakesh Kumar. 2022. “Evolution of Air Pollution Management Policies and Related Research in India.” *Environmental Challenges* 6:100431.

Guptha, Guru Chythanya, Sabyasachi Swain, Nadhir Al-Ansari, Ajay Kumar Taloor, and Deen Dayal. 2022. “Assessing the Role of SuDS in Resilience Enhancement of Urban Drainage System: A Case Study of Gurugram City, India.” *Urban Climate* 41:101075. doi: <https://doi.org/10.1016/j.uclim.2021.101075>.

Hasan, Md Mahmudul, Mirza Mahfuj Hossain, Norizam Sulaiman, and Sayma Khandaker. 2024. “Microsleep Predicting Comparison Between LSTM and ANN Based on the Analysis of Time Series EEG Signal.” *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 16(1):25–31.

Hodson, Timothy O. 2022. “Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not.” *Geoscientific Model Development Discussions* 2022:1–10.

- Horn, Seth A., and Purnendu K. Dasgupta. 2024. "The Air Quality Index (AQI) in Historical and Analytical Perspective a Tutorial Review." *Talanta* 267:125260. doi: <https://doi.org/10.1016/j.talanta.2023.125260>.
- Imrana, Yakubu, Yanping Xiang, Liaqat Ali, and Zaharawu Abdul-Rauf. 2021. "A Bidirectional LSTM Deep Learning Approach for Intrusion Detection." *Expert Systems with Applications* 185:115524. doi: <https://doi.org/10.1016/j.eswa.2021.115524>.
- Indraj, Indraj, and Vishal Warpa. 2024. "Assessment of the Seasonal Trends of Air Pollution: A Case Study of Gurugram City, Haryana, India." *Journal of Air Pollution and Health*.
- Insani, Rachmat Hidayat, and Untung Suropati. n.d. "Comparison of Single Exponential Smoothing and Double Moving Average Algorithms to Forecast Beef Production." *IJID (International Journal on Informatics for Development)* 13(1):448–58.
- Janarthanan, R., P. Partheeban, K. Somasundaram, and P. Navin Elamparithi. 2021. "A Deep Learning Approach for Prediction of Air Quality Index in a Metropolitan City." *Sustainable Cities and Society* 67:102720.
- Kaur, Rajveer, and Puneeta Pandey. 2021. "Air Pollution, Climate Change, and Human Health in Indian Cities: A Brief Review." *Frontiers in Sustainable Cities* 3:705131.
- Kumar Dubey, Ashutosh, Abhishek Kumar, Vicente García-Díaz, Arpit Kumar Sharma, and Kishan Kanhaiya. 2021. "Study and Analysis of SARIMA and

- LSTM in Forecasting Time Series Data.” *Sustainable Energy Technologies and Assessments* 47:101474. doi: <https://doi.org/10.1016/j.seta.2021.101474>.
- Li, Ruoting, Joseph K. Agor, and Osman Y. Özaltın. 2024. “Temporal Pattern Mining for Knowledge Discovery in the Early Prediction of Septic Shock.” *Pattern Recognition* 151:110436.
- Liu, Qian, Bingyan Cui, and Zhen Liu. 2024. “Air Quality Class Prediction Using Machine Learning Methods Based on Monitoring Data and Secondary Modeling.” *Atmosphere* 15(5):553.
- Manikandan, S. Veera, Y. Abilash, S. Hari Prasanth, J. Alfred Daniel, and R. Santhosh. 2023. “Optimized Feature Selection for Air Quality Index Forecasting Using GPR and SARIMA Models.” Pp. 730–35 in *2023 International Conference on Inventive Computation Technologies (ICICT)*. IEEE.
- Minh, Huynh Vuong Thu, Tran Van Ty, Nguyen Dinh Giang Nam, Bui Thi Bich Lien, Nguyen Truong Thanh, Nguyen Phuoc Cong, Gowhar Meraj, Pankaj Kumar, Lam Van Thinh, and Dinh Van Duy. 2024. “Modelling and Predicting Annual Rainfall over the Vietnamese Mekong Delta (VMD) Using SARIMA.” *Discover Geoscience* 2(1):19.
- Natarajan, Suresh Kumar, Prakash Shanmurthy, Daniel Arockiam, Balamurugan Balusamy, and Shitharth Selvarajan. 2024. “Optimized Machine Learning Model for Air Quality Index Prediction in Major Cities in India.” *Scientific Reports* 14(1):6795.

Pandey, Anamika, Michael Brauer, Maureen L. Cropper, Kalpana Balakrishnan, Prashant Mathur, Sagnik Dey, Burak Turkgulu, G. Anil Kumar, Mukesh Khare, Gufran Beig, Tarun Gupta, Rinu P. Krishnankutty, Kate Causey, Aaron J. Cohen, Stuti Bhargava, Ashutosh N. Aggarwal, Anurag Agrawal, Shally Awasthi, Fiona Bennitt, Sadhana Bhagwat, P. Bhanumati, Katrin Burkart, Joy K. Chakma, Thomas C. Chiles, Sourangsu Chowdhury, D. J. Christopher, Subhojit Dey, Samantha Fisher, Barbara Fraumeni, Richard Fuller, Alope G. Ghoshal, Mahaveer J. Golechha, Prakash C. Gupta, Rachita Gupta, Rajeev Gupta, Shreekant Gupta, Sarath Guttikunda, David Hanrahan, Sivadasanpillai Harikrishnan, Panniyammakal Jeemon, Tushar K. Joshi, Rajni Kant, Surya Kant, Tanvir Kaur, Parvaiz A. Koul, Praveen Kumar, Rakesh Kumar, Samantha L. Larson, Rakesh Lodha, Kishore K. Madhipatla, P. A. Mahesh, Ridhima Malhotra, Shunsuke Managi, Keith Martin, Matthews Mathai, Joseph L. Mathew, Ravi Mehrotra, B. V. Murali Mohan, Viswananthan Mohan, Satinath Mukhopadhyay, Parul Mutreja, Nitish Naik, Sanjeev Nair, Jeyaraj D. Pandian, Pallavi Pant, Arokiasamy Perianayagam, Dorairaj Prabhakaran, Poornima Prabhakaran, Goura K. Rath, Shamika Ravi, Ambuj Roy, Yogesh D. Sabde, Sundeep Salvi, Sankar Sambandam, Bhavay Sharma, Meenakshi Sharma, Shweta Sharma, R. S. Sharma, Aakash Shrivastava, Sujeet Singh, Virendra Singh, Rodney Smith, Jeffrey D. Stanaway, Gabrielle Taghian, Nikhil Tandon, J. S. Thakur, Nihal J. Thomas, G. S. Toteja, Chris M. Varghese, Chandra Venkataraman, Krishnan N. Venugopal, Katherine D. Walker, Alexandra Y. Watson, Sarah Wozniak, Denis Xavier, Gautam N. Yadama, Geetika Yadav, D. K. Shukla, Hendrik J. Bekedam, K. Srinath

Reddy, Randeep Guleria, Theo Vos, Stephen S. Lim, Rakhi Dandona, Sunil Kumar, Pushpam Kumar, Philip J. Landrigan, and Lalit Dandona. 2021. “Health and Economic Impact of Air Pollution in the States of India: The Global Burden of Disease Study 2019.” *The Lancet Planetary Health* 5(1):e25–38. doi: [https://doi.org/10.1016/S2542-5196\(20\)30298-9](https://doi.org/10.1016/S2542-5196(20)30298-9).

Perez-Guerra, Uri H., Rassiél Macedo, Yan P. Manrique, Eloy A. Condori, Henry I. Gonzáles, Eliseo Fernández, Natalio Luque, Manuel G. Pérez-Durand, and Manuel García-Herreros. 2023. “Seasonal Autoregressive Integrated Moving Average (SARIMA) Time-Series Model for Milk Production Forecasting in Pasture-Based Dairy Cows in the Andean Highlands.” *Plos One* 18(11):e0288849.

Pramanik, Suvamoy, Carsten Butsch, and Milap Punia. 2021. “Post-Liberal Urban Dynamics in India—The Case of Gurugram, the ‘Millennium City.’” *Remote Sensing Applications: Society and Environment* 22:100504.

Rao, Routhu Srinivasa, Lakshmana Rao Kalabarige, Bhavya Alankar, and Aditya Kumar Sahu. 2024. “Multimodal Imputation-Based Stacked Ensemble for Prediction and Classification of Air Quality Index in Indian Cities.” *Computers and Electrical Engineering* 114:109098. doi: <https://doi.org/10.1016/j.compeleceng.2024.109098>.

Ravindra, Khaiwal, Vikas Singh, and Suman Mor. 2024. “Why We Should Have a Universal Air Quality Index?” *Environment International* 187:108698.

Saturwar, Jitendra, Baljit Thakur, Nayan Siddhapura, and Dilip Sahani. 2022. “Air Quality Index (AQI): Prediction and Optimization.”

Saxena, Aditya, and Binayak Choudhury. 2022. "Internalizing the Externalities of Urban Private Transport – A Case of Gurugram, National Capital Region, India." *Case Studies on Transport Policy* 10(3):1885–97. doi: <https://doi.org/10.1016/j.cstp.2022.08.002>.

Schaidhauer, Ana Carenina Gheller, Fábio Voigt da Costa, and João Carlos Ferreira de Melo-Júnior. 2024. "Air Pollution Generated in an Industrial Region: Effect on the Cardiovascular Health of Humans and Damage Caused to a Plant Species, *Piper Gaudichaudianum* (Piperaceae), Used for Biomonitoring." *Environmental Pollution* 360:124584. doi: <https://doi.org/10.1016/j.envpol.2024.124584>.

Shu, Xiaoling, and Yiwan Ye. 2023. "Knowledge Discovery: Methods from Data Mining and Machine Learning." *Social Science Research* 110:102817.

SHUBHENDER¹, Dr JASPREET HIRA, and PARVEEN BERWAL. n.d. "RETHINKING URBAN MOBILITY IN GURUGRAM: INNOVATIVE SOLUTIONS FOR TACKLING TRAFFIC CONGESTION."

Singh, Damini, Indrani Gupta, and Arjun Roy. 2023. "The Association of Asthma and Air Pollution: Evidence from India." *Economics & Human Biology* 51:101278. doi: <https://doi.org/10.1016/j.ehb.2023.101278>.

Wahid, Abdul, Angga Ade Permana, and Syarif Aminul Khoiri. 2024. "PREDICTION NUMBER OF TOURIST ARRIVALS & PASSENGERS AT THE AIRPORT USE THE TIME MODEL FORCESTING SERIES." Pp. 323–33 in *PROCEEDING INTERNATIONAL CONFERENCE ON*

ECONOMICS, BUSINESS AND INFORMATION TECHNOLOGY (ICEBIT).

Vol. 5.

Yan, Mengyu, and Jinhai Li. 2022. "Knowledge Discovery and Updating under the Evolution of Network Formal Contexts Based on Three-Way Decision."

Information Sciences 601:18–38. doi:

<https://doi.org/10.1016/j.ins.2022.04.010>.

Yan, Rui, Jiaqiang Liao, Jie Yang, Wei Sun, Mingyue Nong, and Feipeng Li. 2021.

"Multi-Hour and Multi-Site Air Quality Index Forecasting in Beijing Using CNN, LSTM, CNN-LSTM, and Spatiotemporal Clustering." *Expert Systems*

with Applications 169:114513. doi:

<https://doi.org/10.1016/j.eswa.2020.114513>.