

BAB II

LANDASAN TEORI

2.1 Polusi Udara

Polusi udara merupakan terkontaminasinya udara dengan zat-zat berbahaya yang dapat menyebabkan dampak negatif terhadap lingkungan, manusia, dan bahkan iklim (Zakiah & Aisyah, 2021). Komponen polusi udara terdiri dari meliputi *partikel matter* 10 (PM_{10}), ozon (O_3), belerang dioksida (SO_2), nitrogen dioksida (NO_2), karbon monoksida (CO) dan *partikel matter* 2,5 ($PM_{2,5}$) (Candrasari et al., 2023). Polusi udara memiliki sumber yang bervariasi dapat bersumber dari alami maupun hasil kegiatan manusia (Zakiah & Aisyah, 2021). Penjelasan setiap komponen polusi yaitu sebagai berikut.

1. *Partikulat Matter* 10 (PM_{10}) dan *Partikulat Matter* 2,5 ($PM_{2,5}$)

Partikulat matter merupakan partikel memiliki ukuran kecil seperti aerosol, debu, dan asap. Partikulat matter 10 (PM_{10}) artinya partikel memiliki diameter kurang dari 10 mikrometer sedangkan partikulat 2,5 ($PM_{2,5}$) merupakan partikel memiliki diameter kurang dari 2,5 mikrometer (Samal et al., 2021a). Kedua partikel tersebut sangat kecil ukurannya sehingga bisa mencapai saluran pernapasan manusia. Partikel ini salah satunya diakibatkan dari emisi kendaraan dan industri (Muthukumar et al., 2022).

2. Ozon (O_3)

Tingkat rendah pada atmosfer yaitu ozon troposferik dapat menyebabkan iritasi pada saluran pernapasan. Ozon (O_3) ini bersumber

emisi industri, emisi kendaraan maupun penguapan bahan kimia (Zhan et al., 2020).

3. Belerang Dioksida (SO_2)

Belerang dioksida berwujud gas dapat menyebabkan tidak hanya infeksi saluran pernapasan tetapi berkontribusi dalam pembentukan hujan asam. Penyebab pembentukan belerang dioksida dari pembakaran batu bara, pembakaran minyak dan proses peleburan logam (Candrasari et al., 2023).

4. Nitrogen Dioksida (NO_2)

Nitrogen dioksida tidak hanya menyebabkan iritasi pada saluran pernapasan manusia tetapi menjadi penyebab pembentukan ozon troposferik (H. Liu et al., 2021). Nitrogen dioksida disebabkan pembakaran bahan bakar fosil.

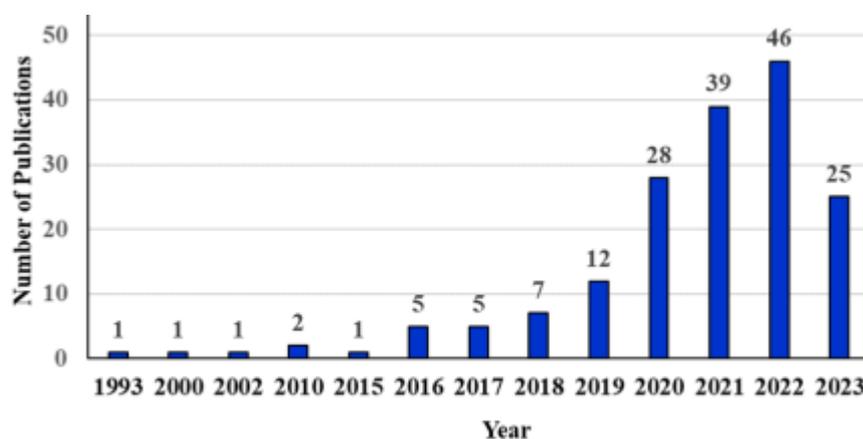
5. Karbon Monoksida (CO)

Karbon monoksida dapat mengikat hemoglobin di dalam darah sehingga dapat menghambat laju dari oksigen dan menyebabkan hipoksia. Penyebab pembentukan karbon monoksida terjadi karena pembakaran yang tidak sempurna pada proses pembakaran bahan bakar dalam kendaraan maupun industri (Tsokov et al., 2022).

Polutan partikel matter 10 (PM_{10}), ozon (O_3), belerang dioksida (SO_2), nitrogen dioksida (NO_2), karbon monoksida (CO) dan partikel matter 2,5 ($PM_{2,5}$) menjadi parameter di dalam perhitungan *Air Quality Index* (AQI) atau indeks kualitas udara (Seng et al., 2021). Skala numerik untuk mengukur dan mengategorikan kualitas udara dalam suatu periode tertentu yaitu menggunakan

Air Quality Index (AQI). Pengukuran setiap polutan dinormalisasi dan dihitung dalam indeks tunggal untuk menggambarkan kondisi kualitas udara secara keseluruhan.

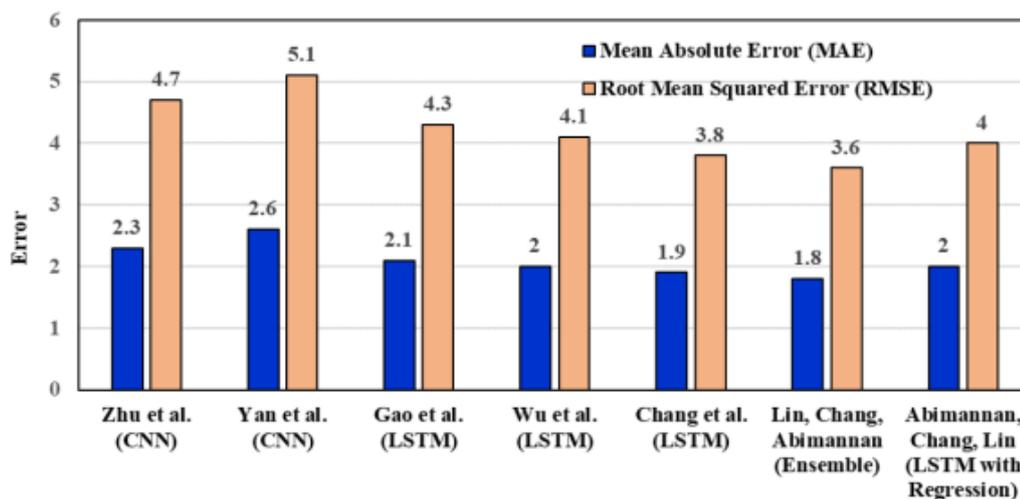
Penggunaan teknik *deep learning* memiliki potensi besar dalam memprediksi kualitas udara. berbagai arsitektur jaringan termasuk, *convolutional neural networks*, *recurrent neural networks* dan *Long Short Term Memory* sebagai pendekatan dalam peramalan kualitas udara (Liao et al., 2020). Berikut gambar 2.1 banyaknya penelitian prediksi kualitas udara tahunan yang menggunakan *deep learning*.



Gambar 2. 1 Penelitian prediksi udara (Abimannan et al., 2023)

Penelitian mengenai prediksi kualitas udara menggunakan *deep learning* dapat dilihat pada gambar 2.1. Terlihat terjadi jumlah kenaikan yang signifikan dari 2019 yang berjumlah 12 publikasi hingga tahun 2022 menjadi tahun terbanyak penelitian sebanyak 46 publikasi. Selain itu pada tahun 2023 mengalami penurunan hanya terdapat 25 publikasi. Berikut gambar 2.2

perbandingan dari performa model *deep learning* yang telah dilakukan dalam penelitian sebelumnya.



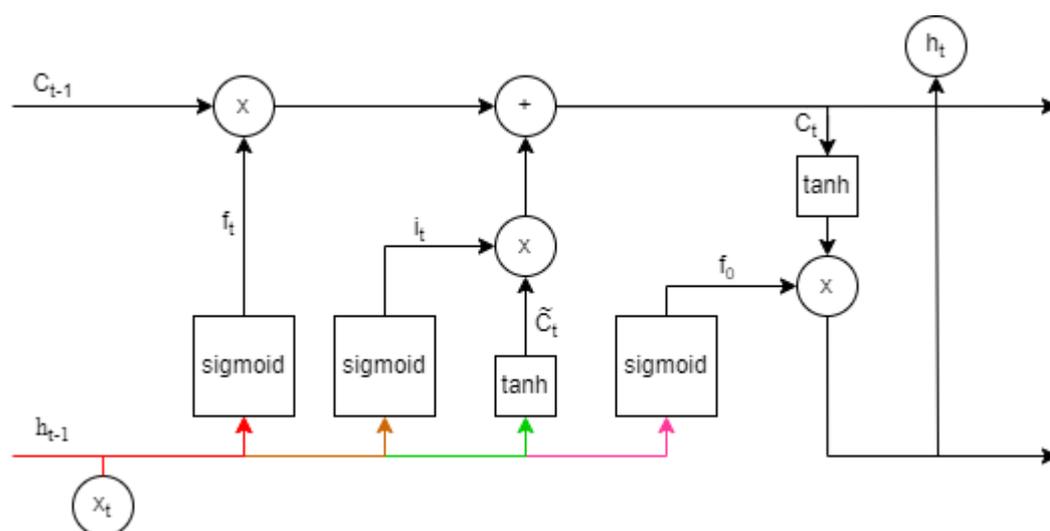
Gambar 2. 2 Penelitian prediksi udara (Abimannan et al., 2023)

Pendekatan *deep learning* yang dilakukan untuk memprediksi kualitas udara dapat dilihat pada gambar 2.2. Terlihat CNN memiliki nilai tertinggi RMSE 5.1 dan MAE 2.6. Nilai terendahnya yaitu Model LSTM dengan nilai RMSE 3.8 dan MAE 1.9, serta model ensemble RMSE 3.6 dan MAE 1.8. LSTM dan ensemble memiliki perbedaan hanya terlampau sedikit dan lebih baik pada ensemble.

2.2 BiLSTM

BiLSTM atau *Bidirectional Long Short-Term Memory* merupakan salah satu dari arsitektur *Recurrent Neural Network* (RNN) untuk memproses data berurutan secara efektif. Komponen BiLSTM terdapat *Long Short Term-Memory* (LSTM) memiliki kemampuan dalam menyimpan informasi penting dan mengabaikan informasi yang tidak diperlukan (Feng & Chen, 2021). LSTM dapat

mempelajari ketergantungan jangka panjang di dalam urutan data sehingga memungkinkan dalam memahami pola kompleks selama interval lebih panjang (Espinosa et al., 2021). Model LSTM memiliki tiga kunci utama yaitu, *forget gate*, *input gate*, dan *output gate* untuk mengontrol informasi masuk dan keluar dari sel LSTM (Cheng et al., 2021). Arsitektur dari LSTM dapat dilihat pada gambar 2.3.



Gambar 2. 3 Model LSTM (Drewil & Al-Bahadili, 2022).

Pada gambar 2.3 merupakan gambar model dari LSTM. Model LSTM memiliki satu *cell state* dan tiga *gate*. Setiap *gate* dan *cell state* memiliki fungsi dan perhitungan yang berbeda. Berikut penjelasan dan perhitungan di setiap *gate* dan *cell state*.

1. Forget Gate

Forget gate merupakan bagian dari model LSTM yang berfungsi untuk menentukan seberapa pentingnya suatu informasi di dalam memori sel LSTM

(Wang, Li, et al., 2022). Pada gambar 2.3, terlihat adanya garis berwarna merah yang mewakili *forget gate*. *Forget gate* berfungsi untuk memproses informasi dari dua sumber, yaitu hidden state sebelumnya (h_{t-1}) dan input pada waktu (x_t). Jika dilihat dari persamaan (1) maka sebagai berikut (D. Li et al., 2022).

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

Persamaan (1) *forget gate* (f_t) pada waktu t ditentukan menggunakan fungsi aktivasi sigmoid (σ), yang mengubah hasil operasi linear dari matriks bobot (W_f) dan vektor gabungan $[h_{t-1}, x_t]$. Vektor gabungan ini berisi *hidden state* sebelumnya (h_{t-1}) dan input (x_t). Selain itu, vektor bias (b_f) ditambahkan ke hasil operasi untuk memberikan penyesuaian tambahan. Fungsi sigmoid digunakan untuk menghasilkan bobot bernilai antara 0 dan 1. Jika nilai mendekati 0 maka informasi akan dilupakan, sementara itu jika nilai mendekati 1 maka informasi akan disimpan.

2. *Input Gate*

Input gate berfungsi untuk menentukan informasi yang akan disimpan di *cell state* (B. Liu et al., 2021). Pada gambar 2.3 terdapat garis berwarna oranye dan hijau yang mewakili *input gate* pada LSTM. Berikut persamaan (2) untuk garis berwarna oranye pada model LSTM sebagai berikut (D. Li et al., 2022).

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

Persamaan (2) di atas, i_t merupakan output dari fungsi sigmoid untuk menentukan banyaknya informasi baru untuk disimpan di *cell state*. Nilai i_t pada

langkah waktu t dihasilkan dari perkalian matriks bobot W_i dengan vektor gabungan ($[h_{t-1}, x_t]$) dan menambahkan vektor bias (b_i). Setelah itu, pada proses input juga terdapat persamaan (3) perhitungan untuk garis berwarna hijau pada model LSTM sebagai berikut (D. Li et al., 2022).

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

Input gate terdapat proses perhitungan untuk mengetahui nilai *candidate state* (\tilde{C}_t). Persamaan (3) di atas menyatakan bahwa \tilde{C}_t pada langkah waktu t dihasilkan dari operasi tangen hiperbolik terhadap hasil penjumlahan dari perkalian matriks bobot W_c dengan vektor gabungan $[h_{t-1}, x_t]$ dan menambahkan vektor bias (b_c). Nilai \tilde{C}_t akan bernilai rentang antara -1 dan 1 dikarenakan menggunakan fungsi aktivasi tangen hiperbolik.

3. *Cell State*

Cell state pada LSTM merupakan salah satu komponen kunci karena *cell state* berfungsi untuk menyimpan informasi jangka panjang yang dapat digunakan oleh model (Huang et al., 2022). Persamaan (4) untuk penyimpanan informasi di dalam *cell state* sebagai berikut (D. Li et al., 2022).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

Persamaan (4) merupakan proses dari menentukan nilai *cell state* (C_t). Nilai C_t dihasilkan dari perkalian *forget gate* (f_t) dan *cell state* sebelumnya (C_{t-1}) kemudian, ditambahkan dengan hasil perkalian *input gate* (i_t) dan *candidate state* (\tilde{C}_t). Persamaan tersebut menunjukkan bahwa LSTM mampu untuk mengelola

informasi jangka panjang serta dapat mempertahankan informasi atau mengabaikan informasi yang tidak penting.

4. *Output gate*

Output gate berfungsi dalam mengatur sebanyak output yang akan dijadikan hidden state dengan perkalian antar elemen. Pada gambar 2.3 terdapat garis berwarna pink yang menunjukkan *output gate* (Wang, Li, et al., 2022). Secara matematis, persamaan (5) *output gate* pada LSTM dihitung menggunakan nilai sebagai berikut (D. Li et al., 2022).

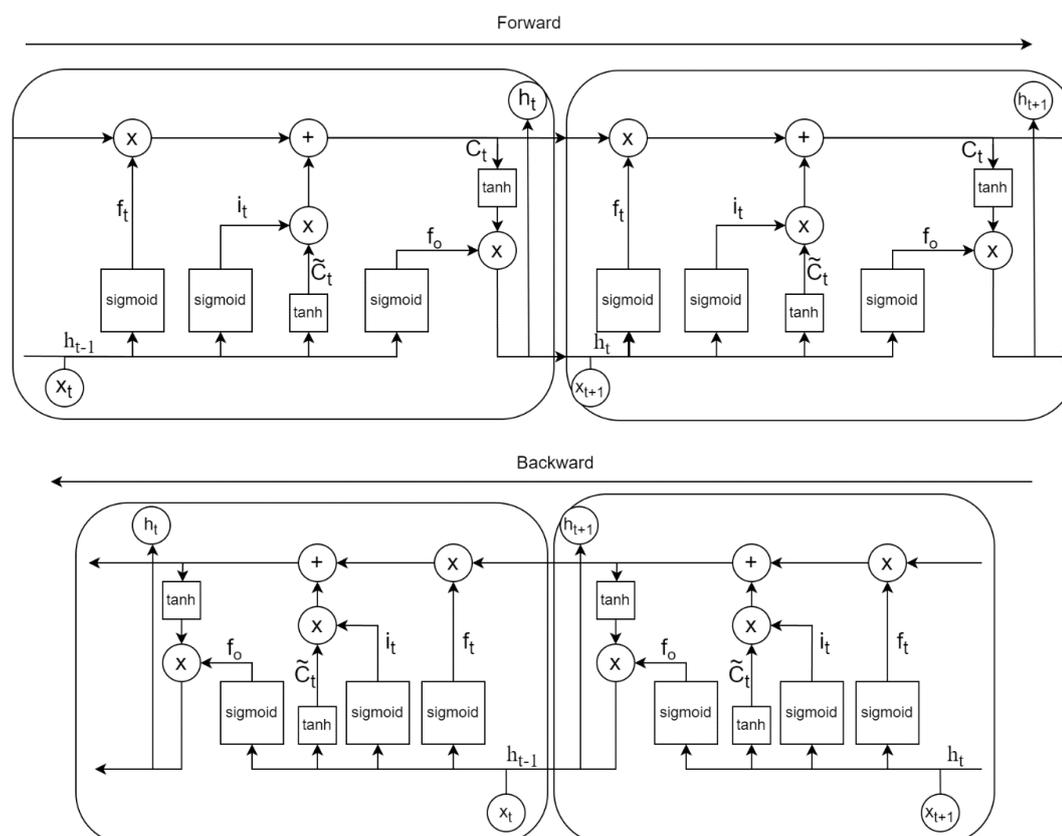
$$f_o = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

Persamaan (5) menghitung *output gate* (f_o) menggunakan matriks bobot (W_o), vektor *input* gabungan ($[h_{t-1}, x_t]$) dan vektor bias (b_o). Fungsi sigmoid memastikan bahwa f_o berada dalam rentang 0 hingga 1 sesuai dengan interpretasi probabilitas. Kemudian, nilai dari f_o digunakan untuk mencari nilai *hidden state* sebagai *output*. Persamaan (6) *hidden state* pada LSTM sebagai berikut (D. Li et al., 2022).

$$h_t = f_o \cdot \tanh(C_t) \quad (6)$$

Persamaan (6) *hidden state* (h_t) membutuhkan nilai *output gate* (f_o) dan *cell state* (C_t) dalam fungsi tangen hiperbolik untuk menghasilkan nilai antara -1 dan 1. Nilai *hidden state* didapatkan dari perkalian f_o dan $\tanh(C_t)$ untuk mengatur banyaknya *cell state* yang diungkapkan sebagai *output*.

Bidirectional Long Short-Term Memory memanfaatkan model LSTM yang berjalan dua arah yaitu *forward* dan *backward* (Aamir et al., 2022). Kombinasi kedua arah membuat model dapat memahami urutan yang kompleks dari kedua arah (Aamir et al., 2022). Model pada BiLSTM dapat dilihat pada gambar 2.4.



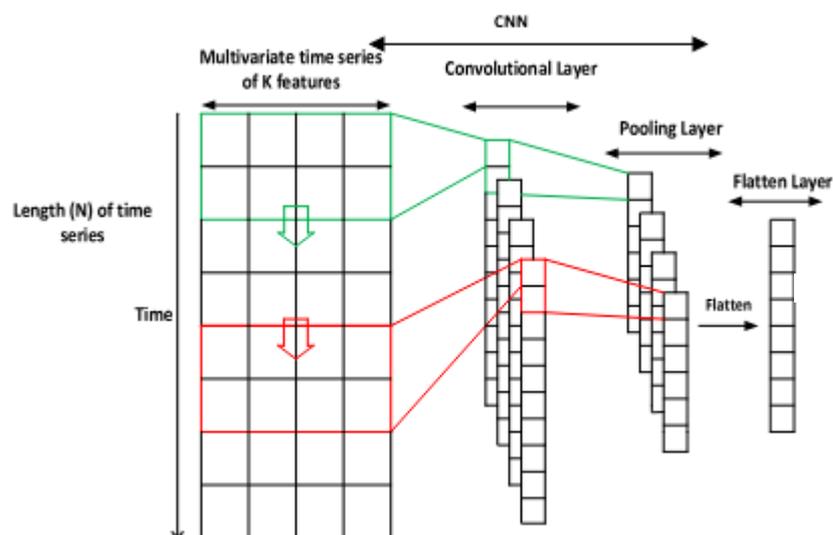
Gambar 2. 4 Model BiLSTM (Zhang et al., 2023).

BiLSTM memiliki kesamaan dengan LSTM, sebagaimana terlihat pada gambar 2.4. Prinsip perhitungannya sama dengan LSTM, namun BiLSTM berbeda dalam hal prosesnya yang dilakukan secara dua arah, yaitu *forward* dan *backward* (Zhang et al., 2023). Pada langkah waktu pertama, input diberikan untuk langkah *forward*, sementara pada langkah waktu terakhir, input diberikan

untuk langkah *backward* (Aamir et al., 2022). *Output* dari kedua arah tersebut pada setiap langkah waktu digabungkan, memungkinkan model untuk memiliki pemahaman yang lebih holistik dan komprehensif.

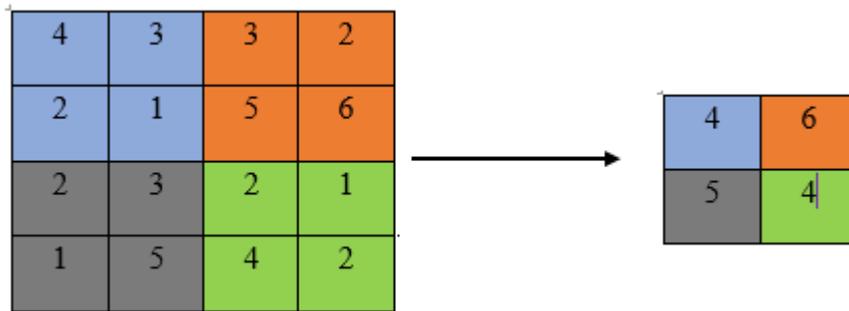
2.3 Convolutional Neural Network

Convolutional Neural Network (CNN) merupakan salah satu jenis *neural network* untuk pengenalan pola seperti gambar, video dan data historis. CNN bisa difokuskan pada data satu dimensi digunakan pada tugas data urutan waktu atau deret, seperti data historis pada *Air Quality Index* (AQI) (Wang, Li, et al., 2022). Pada layer satu dimensi memfilter data yang bergerak sepanjang dimensi data input. Filter digunakan dalam ekstraksi fitur dari urutan data pada setiap langkah, filter akan melakukan operasi konvolusi dari data input (Wang, Jin, et al., 2022). Berikut gambar 2.5 arsitektur dari *convolutional neural network*.



Gambar 2. 5 Arsitektur CNN (Lawal et al., 2021).

Arsitektur dari CNN dapat dilihat pada gambar 2.5. Arsitektur CNN tersebut khusus untuk menerima input data satu dimensi. CNN terdiri dari lapisan konvolusi satu dimensi, *pooling layer*, *flatten layer* atau *ouput layer*. Beberapa hal harus di perhatikan di dalam *convolution layer* sebagai berikut. *Input* data dapat berupa *time series* dengan lebar layer konvolusi sama dengan deret waktu dan lebar deret waktu sama dengan jumlah fitur pada K dan panjang deret N . *Pooling layer* dari setiap vektor deret waktu dari lapisan konvolusi membentuk vektor baru. Kemudian, dari *pooling layer* diteruskan ke *flatten layer* ataupun *layer output* yang disesuaikan dengan kebutuhan model (Lawal et al., 2021). Proses *pooling layer* biasanya menggunakan teknik *max pooling layer* seperti pada gambar 2.6



Gambar 2. 6 *Pooling Layer*

Gambar 2.6 merupakan dari contoh proses *pooling layer* pada CNN. Proses *pooling layer* yaitu memiliki sebuah filter dengan ukuran *kernel size* dan *stride* di dalam memfilter *feature map* (Alzubaidi et al., 2021). Gambar 2.5 menunjukkan bahwa dengan *max pool 2 x 2* dan *stride* dua dari *feature map* yang

sebelah kiri diambil nilai max sehingga menghasilkan ukuran 2×2 *feature map* sebelah kanan.

2.4 Penelitian Terkait

Terdapat beberapa penelitian yang signifikan menggunakan berbagai pendekatan untuk memodelkan prediksi kualitas udara. Kontribusi penting dari penelitian-penelitian tersebut terletak pada pemahaman dan prediksi faktor-faktor yang mempengaruhi kualitas udara. Temuan-temuan ini memiliki potensi untuk menjadi dasar penting dalam pengembangan metode prediksi yang lebih tepat dan efisien dalam pemantauan kualitas udara. Berikut Tabel 2.1 penelitian terkait menjadi landasan dalam penelitian ini.

Tabel 2. 1 Penelitian Terkait

Penulis	Algoritma	Fitur	Fokus Pembahasan
(Mengara Mengara et al., 2022)	<i>Attention-based Convolutional BiLSTM</i>	PM_{10} dan $PM_{2.5}$	Model mampu memperhitungkan kolerasi intriks antara polusi partikel di lokasi yang berbeda. Hasil evaluasi model menunjukkan bahwa model memiliki performa yang baik untuk memprediksi jangka pendek atau panjang. Model memiliki kompleksitas tinggi karena

			melibatkan gabungan beberapa model. Penggunaan data paralelisme memerlukan komputasi yang tinggi.
(Jinsong Zhang, Yongtao Peng, 2019)	CNN+BILSTM+ <i>Attention</i>	$PM_{2.5}$	Pendekatan <i>hybrid</i> dengan menggabungkan keunggulan setiap modelnya. CNN mengekstrak fitur spasial, BiLSTM menangkap pola temporal, dan attention memberik bobot informasi yang relevan. Fungsi aktivasi, <i>batch size</i> , <i>dropout rate</i> , dan <i>learning rate</i> mungkin perlu lebih rinci untuk memudahkan reproduksi dan penyesuaian model oleh peneliti lain.
(Rani Samal et al., 2020)	CNN+LSTM+ <i>Support Vector Regressor</i>	PM_{10}	Model mampu memberikan prediksi kualitas udara untuk dua minggu ke depan dengan tingkat akurasi 91-96% dari pada model dasar.

(Chowdhury et al., 2020)	LSTM	AQI	LSTM dapat menangani data deret waktu baik. Memodelkan hubungan jangka panjang dalam data. Model memerlukan volume data yang besar dalam <i>training</i> untuk mengoptimalkan performa.
(Ma et al., 2020)	LSTM based Bayesian Optimization	$PM_{2.5}$	Model ini memiliki kemampuan untuk mengoptimal parameter model dalam meningkatkan kinerja prediksi dengan menemukan kombinasi parameter yang optimal.
(Aamir et al., 2022)	BiLSTM	CO	Model lebih baik daripada LSTM dan GRU. Hasil analisis BiLSTM menghasilkan nilai MAE, MSE, dan MAPE lebih rendah. BiLSTM membutuhkan komputasi yang besar dibandingkan

			model LSTM dan GRU.
(H. Liu et al., 2021)	CNN+LSTM	$PM_{2.5}$	Pemodelan kombinasi CNN dan LSTM menghasilkan nilai RMSE = 14.3 dalam memprediksi $PM_{2.5}$.
(Huang et al., 2022)	<i>Dual Attentions-Gated Recurrent Unit with Skip Connections</i>	AQI	Dual <i>attention</i> fokus pada bagian penting dari input, dapat memperbaiki kinerja dengan memberikan bobot berbeda pada bagian urutan waktu. GRU dapat mengatasi <i>vanishing gradient</i> . Skip connection memungkinkan informasi dari layer <i>input</i> langsung ke layer <i>output</i> .
(Samal et al., 2021a)	<i>Convolutional</i> + LSTM	$PM_{2.5}$	Model dapat mengidentifikasi fitur tersembunyi dalam <i>dataset</i> yang besar. Model Mengakibatkan waktu pelatihan yang lama dan kebutuhan komputasi yang lebih tinggi.
(L. Liu &	LSTM- <i>Staked</i>	AQI	Model LSTM berfungsi

Zhang, 2021)	<i>Auto-Encoder</i>		dalam evaluasi prediksi polusi udara dan dikombinasikan dengan <i>stacked auto encoder</i> sehingga model dapat mengestrak komponen intrinsik pada data.
(Gupta et al., 2023)	<i>Random forest Regression</i>	$PM_{2.5}$, CO , NO_2 , O_3 , SO_2 , NH_3	Model <i>random forest regression</i> membangun <i>multiple decision trees</i> dengan mengambil rata-rata prediksi sehingga model lebih general dan dapat meminimalkan efek <i>overfitting</i> . Model menghasilkan MSE = 0.1464, RMSE = 0.3826 dan MAE = 0.2163
(Drewil & Al-Bahadili, 2022)	<i>Genetic Algorithm+LSTM</i>	$PM_{2.5}$, PM_{10} , CO , NO_3	Penggunaan <i>Genetic Algorithm</i> membantu LSTM dalam menembukan <i>hyperparameter</i> secara efisien dalam meningkatkan

			performa model LSTM.
(Arkadia et al., 2022)	LSTM	$PM_{10}, SO_2, CO, O_3, NO_3$	Model memproses data time series dan dioptimasi dengan adam optimizer dalam memastikan setiap layer memberikan nilai yang optimal. Hasil prediksi yaitu nilai MAPE pada $PM_{10} = 4.37, SO_2 = 5.02, CO = 18.50, O_3 = 5.23$ dan $NO_3 = 37.28$
(Khan et al., 2022)	CNN+LSTM	$PM_{10}, SO_2, CO, O_3, NO_2, PM_{2.5}$	Model CNN-LSTM mampu mengekstraksi fitur yang efisien dan mampu memahami data historis. Dilakukan analisis temporal untuk $PM_{10}, SO_2, CO, O_3, NO_2,$ dan $PM_{2.5}$
(B. Liu et al., 2021)	<i>Discrete Wavelet Transform</i> +LSTM	$PM_{10}, SO_2, CO, O_3, NO_2, PM_{2.5}$	Model mampu meningkatkan akurasi prediksi NO_2 karena discrete wavelet transform dapat mendekomposisi data input untuk membantu dalam mengidentifikasi fitur-fitur

			sehingga meningkatkan pemahaman pola dan tren.
(Seng et al., 2021)	LSTM	$SO_2, O_3, PM_{2.5}, CO, NO_2,$	Model dapat memprediksi beberapa indikator kualitas udara yaitu, $SO_2, CO, O_3, NO_2,$ dan $PM_{2.5}$ untuk mengastasi berbagai aspek polusi secara bersamaan. Kemampuan <i>multioutput</i> meningkatkan kegunaan model
(Aliyudin et al., 2024)	BiLSTM+CNN	AQI, $PM_{10}, SO_2, CO, PM_{2.5}, O_3, NO_2$	Pengintegrasian Layer CNN dengan BiLSTM untuk memprediksi AQI, $PM_{10}, SO_2, CO, PM_{2.5}, O_3, NO_2.$ Pada pengembangan model layer CNN untuk ekstraksi spasial pada data input dan Layer BiLSTM untuk ekstraksi sekuensial.

Pada tabel 2.1, berbagai penelitian dalam prediksi kualitas udara telah mengusulkan berbagai model. Penelitian oleh (Mengara Mengara et al., 2022) dan (Jinsong Zhang, Yongtao Peng, 2019) menggunakan model berbasis *attention*,

CNN, dan BiLSTM, kemudian penelitian oleh (Rani Samal et al., 2020) menggunakan CNN, LSTM dan SVR, dari beberapa penelitian tersebut menunjukkan performa baik dalam memprediksi polusi udara ditunjukkan dengan akurasi yang signifikan walaupun memiliki kompleksitas yang tinggi. Penelitian yang dilakukan oleh (Chowdhury et al., 2020), (Ma et al., 2020), (Aamir et al., 2022), (Arkadia et al., 2022), dan (Seng et al., 2021), memiliki model yang digunakan cukup sederhana hanya menggunakan LSTM atau BiLSTM. Kombinasi sederhana seperti pada penelitian oleh (H. Liu et al., 2021), (Samal et al., 2021a), dan (Khan et al., 2022), menggunakan CNN dikombinasikan dengan LSTM tetapi tidak difokuskan untuk memprediksi seluruh komponen kualitas udara. Penggunaan model LSTM yang dilakukan oleh (Drewil & Al-Bahadili, 2022) dengan menggunakan *Genetic Algorithm* untuk menentukan *hyperparameter* untuk membuat model paling efektif. Kombinasi GRU dengan *attentions* dilakukan dalam penelitian (Huang et al., 2022), model difokuskan untuk memperbaiki kinerja bobot pada saat *training*. Namun, beberapa penelitian belum terdapat yang secara khusus mengintegrasikan model CNN dengan BiLSTM kemudian mempertimbangkan seluruh fitur komponen polusi udara untuk prediksi kualitas udara.

Oleh karena itu, penelitian ini merupakan salah satu upaya untuk mengisi kesenjangan yang masi ada dalam model prediksi kualitas udara. Mengintegrasikan layer CNN dengan BiLSTM untuk mengekstraksi pola spasial dan sekuensial dalam upaya meningkatkan performa model. Selain itu, model tidak

hanya menggunakan satu komponen polusi udara tetapi semua komponen polusi udara untuk menambah pemahaman terhadap model dalam memprediksi.

Perbandingan penelitian terkait dapat dilihat melalui matriks penelitian pada tabel 2.2.

Tabel 2. 2 Matrik Penelitian

Penulis	Algoritma	Fitur						
		<i>AQI</i>	<i>PM₁₀</i>	<i>SO₂</i>	<i>NO₂</i>	<i>CO</i>	<i>O₃</i>	<i>PM_{2.5}</i>
(Mengara Mengara et al., 2022)	<i>Attention-based Convolutional BiLSTM</i>	-	√	-	-	-	-	√
(Jinsong Zhang, Yongtao Peng, 2019)	<i>CNN+BiLSTM+ Attention</i>	-	-	-	-	-	-	√
(Rani Samal et al., 2020)	<i>CNN+LSTM+ Support Vector Regressor</i>	-	√	-	-	-	-	-
(Chowdhury et al., 2020)	<i>LSTM</i>	√	-	-	-	-	-	-
(Ma et al., 2020)	<i>LSTM based Bayesian Optimization</i>	-	-	-	-	-	-	√

(Aamir et al., 2022)	BiLSTM	-	-	-	-	√	-	-
(H. Liu et al., 2021)	CNN+LSTM	-	-	-	-	-	-	√
(Huang et al., 2022)	<i>Dual Attentions-Gated Recurrent Unit with Skip Connections</i>	√	-	-	-	-	-	-
(Samal et al., 2021a)	<i>Convolutional + LSTM</i>	-	-	-	-	-	-	√
(L. Liu & Zhang, 2021)	<i>LSTM- Staked Auto-Encoder</i>	√	-	-	-	-	-	-
(Gupta et al., 2023)	<i>Random forest Regression</i>	√	-	√	√	√	√	√
(Drewil & Al-Bahadili, 2022)	<i>Genetic Algorithm+LSTM</i>	-	√	-	√	√	-	√
(Arkadia et al., 2022)	LSTM	-	√	√	√	√	√	-
(Khan et al., 2022)	CNN+LSTM	-	√	√	√	√	√	√
(B. Liu et al., 2022)	<i>Discrete Wavelet</i>	-	√	√	√	√	√	√

al., 2021)	<i>Transform</i> +LSTM							
(Seng et al., 2021)	LSTM	-	-	√	√	√	√	√
(Aliyudin et al., 2024)	BiLSTM+CNN	√	√	√	√	√	√	√

Pada tabel 2.2, matrik penelitian menunjukkan perbedaan diantara penelitian terkait lainnya. Penelitian sebelumnya telah melakukan berbagai kombinasi prediksi kualitas udara, seperti LSTM, CNN, BiLSTM dan kombinasi lainnya. Penelitian mengkombinasikan model cukup kompleks seperti yang dilakukan (Mengara Mengara et al., 2022) dan (Jinsong Zhang, Yongtao Peng, 2019) menggunakan *Attention-Based Convolutional BiLSTM*. Kemudian, penelitian sebelumnya sering kali terbatas pada pasangan algoritma yang lebih sederhana, seperti pada penelitian (Khan et al., 2022), (Samal et al., 2021a), dan (H. Liu et al., 2021) menggunakan kombinasi CNN dengan LSTM. Pada penelitian sebelumnya model cenderung difokuskan pada satu atau beberapa polutan komponen polusi udara tanpa mengintegrasikan semua parameter komponen polusi udara. Penelitian yang dilakukan oleh (Aliyudin et al., 2024) berkontribusi pada beberapa kesenjangan pada penelitian-penelitian sebelumnya. Penelitian ini menggunakan kombinasi BiLSTM dan CNN yang belum banyak dieksplorasi secara mendalam. Kemudian, penelitian ini menyertakan semua parameter komponen polusi udara dalam model prediksi. Mencakup seluruh

parameter, model mampu memberikan gambaran yang lebih menyeluruh tentang kondisi polusi udara. Pada penelitian yang dilakukan oleh (H. Liu et al., 2021) dan (Mengara Mengara et al., 2022), data *time series* mengenai kondisi polutan tertentu digunakan dengan fokus pada suatu wilayah sebagai objek penelitian. Demikian juga, penelitian yang dilakukan oleh (Aliyudin et al. 2024) memilih wilayah yang sering mengalami polusi udara, dengan menggunakan dan menganalisis data *time series* indeks kualitas udara serta polutannya untuk kebutuhan pengembangan model.