

BAB I

PENDAHULUAN

1.1 Latar Belakang

Polusi udara memberikan dampak buruk pada lingkungan dan kesehatan manusia. Polusi udara menyebabkan berbagai masalah kesehatan seperti, kanker paru-paru, asma, dan penyakit pernapasan kronis (Dina *et al.*, 2019). Jakarta merupakan sebuah kota yang menjadi pusat ekonomi bisnis di Indonesia, sehingga menjadikannya salah satu kota dengan polusi tertinggi di Indonesia (Nursyarifah, 2022). Berdasarkan data dinas kesehatan di Jakarta tahun 2023 yakni periode januari sampai juni terdapat 638.291 orang mengalami infeksi saluran pernapasan (ISPA). Meningkatnya konsentrasi polutan seperti $PM_{2.5}$ dan NO_2 , masyarakat Jakarta akan lebih rentan terhadap gangguan kesehatan pernafasan dan tingkat tinggi polutan CO dapat merusak keseimbangan ekosistem (Rendana *et al.*, 2022). Terjadinya polusi udara disebabkan oleh kegiatan industri, minim ruang terbuka hijau, dan asap transportasi (Zakiah & Aisyah, 2021). Komponen polusi udara 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). Oleh sebab itu pentingnya pemantauan kualitas udara dengan memprediksi kualitas udara beserta komponen polusi udara untuk memberikan peringatan secara dini dalam upaya menjaga kesehatan.

Pendekatan menggunakan *deep learning* dalam memprediksi kualitas udara beserta komponennya polusi udara telah dilakukan untuk menjaga kesehatan manusia dari polusi udara (Abimannan et al., 2023). Data polusi udara bersifat *time series* yaitu mencakup pengamatan kualitas udara beserta komponennya dalam interval waktu tertentu (Tsokov et al., 2022). Penelitian oleh (Janarthanan et al., 2021) dan (Seng et al., 2021), penggunaan model *multi-output* dan *multi-index of supervised learning* (MMSL) berbasis *Long Short-Term Memory* (LSTM) dalam memprediksi indikator polusi udara $PM_{2.5}$, CO , NO_2 , O_3 dan SO_2 , model tersebut dapat memahami beragam data dan menangkap dependensi jangka panjang dalam data. Perbandingan model *Bidirectional Long Short Term Memory* (BiLSTM), LSTM, *Convolutional Neural Network* (CNN), dan *Back Propagation Neural Network* (BPNN) dalam memprediksi kualitas udara untuk mengetahui model yang paling optimal telah dilakukan oleh (Zhan et al., 2020) kemudian hasilnya, model BiLSTM menghasilkan performa terbaik yaitu nilai *Root Mean Square Error* (RMSE) 22,60, *Mean Absolute Error* (MAE) 15,87, dan *Mean Square Error* (MSE) 510,85. BiLSTM menunjukkan performa terbaik dibandingkan dengan *Gated Recurrent Unit* (GRU) dan *Temporal Convolutional Network* (TCN), dan terjadi peningkatan kinerja saat diintegrasikan dengan *Light Gradient Boosting Machine* (LGBM) (Zhang et al., 2023).

Model BiLSTM memiliki kemampuan untuk memahami data *time series* secara menyeluruh, namun kurang efektif dalam mengidentifikasi dan memahami pola tren naik atau turun kualitas udara secara spasial dalam rentang

waktu tertentu di dalam data *time series* (Riyadi & Jasmir, 2023). Model CNN memiliki *convolutional layer* yang dirancang untuk memahami atau mengekstraksi aspek spasial dari data *time series* terkait polusi udara (Tsokov et al., 2022). Hasil kajian pada *survey paper* menunjukkan, CNN memiliki performa yang kurang optimal dengan nilai MAE dan RMSE yang tinggi, dikarenakan model hanya mampu memahami data secara spasial pada data *time series* (Abimannan et al., 2023). Penggabungan model *Variational Mode Decomposition* (VMD) dan LSTM menghasilkan kinerja yang lebih baik, terutama ketika model tersebut diintegrasikan dengan CNN (Han et al., 2023). Kinerja model LSTM meningkat saat diubah menjadi BiLSTM, dan efisiensinya dapat ditingkatkan lebih lanjut dengan diintegrasikan CNN, sebagaimana terbukti oleh penurunan nilai RMSE dan MAE (D. Li et al., 2022). Penerapan CNN dapat memberikan kemudahan mengekstraksi data secara spasial dalam data *time series* konsentrasi polutan udara (Rani Samal et al., 2020).

Berdasarkan seluruh uraian persoalan yang telah dibahas tersebut, terdapat kesenjangan penelitian yang masih tersisa, yaitu model BiLSTM memiliki kekurangan dalam ekstraksi pola spasial pada data *time series*. Kemudian, model CNN merupakan model kurang optimal dibandingkan dengan model lainnya dikarenakan pada kinerjanya CNN hanya ekstraksi pola spasial pada data *time series*. Sehingga dalam penelitian ini melakukan pengintegrasian model BiLSTM dengan layer *convolutional* untuk memprediksi kualitas udara beserta komponennya.

Peluang penelitian ini yaitu mengembangkan model BiLSTM dengan

mengintegrasikan *convolutional* layer yang diintegrasikan untuk menciptakan model yang lebih baik dalam memprediksi kualitas udara. Model pengembangan tersebut dapat mengekstraksi pola spasial maupun sekuensial secara keseluruhan sehingga diharapkan menghasilkan model yang lebih optimal.

1.2 Rumusan Masalah

Penelitian ini berfokus pada rumusan masalah yang dispesifikasikan dalam bentuk pertanyaan sebagai berikut:

1. Bagaimana menerapkan model BiLSTM yang diintegrasikan dengan penambahan layer *convolutional* untuk kebutuhan ekstraksi pola spasial dan sekuensial pada data *time series*?
2. Bagaimana pengaruh pengembangan model BiLSTM diintegrasikan dengan layer *convolutional* terhadap prediksi indeks kualitas udara?

1.3 Tujuan Penelitian

Penelitian ini memiliki tujuan yang ditentukan secara spesifik sesuai dengan rumusan masalah yang telah dirumuskan sebagai berikut:

1. Mengintegrasikan model BiLSTM dengan layer *convolutional* untuk kebutuhan ekstraksi pola spasial dan sekuensial dalam pemodelan memprediksi kualitas udara.
2. Mengukur pengaruh pengintegrasian layer *convolutional* pada BiLSTM dalam memprediksi kualitas udara.

1.4 Manfaat Penelitian

Manfaat penelitian ini ditentukan secara spesifik menjadi beberapa poin sebagai berikut:

1. Pemodelan BiLSTM dikembangkan dengan penambahan layer *convolutional* sehingga model dapat mengekstraksi pola spasial maupun sekuensial untuk menghasilkan model yang lebih optimal.
2. Model BiLSTM diintegrasikan dengan layer *convolutional* dapat memprediksi tidak hanya indeks kualitas udara tetapi beserta komponen udara lainnya.
3. Hasil prediksi tersebut memberikan informasi keadaan udara dalam upaya meningkatkan kesadaran manusia terhadap kondisi udara untuk menjaga kesehatan.

1.5 Batasan Masalah

Penelitian ini menetapkan batasan-batasan masalah supaya mengarahkan penelitian ke arah yang lebih spesifik. Berikut adalah rincian batasan-batasan yang diterapkan dalam penelitian ini:

1. Penelitian ini berfokus pada pengembangan model BiLSTM dengan penambahan layer *convolutional* untuk memprediksi *multioutput* yaitu indeks kualitas udara, PM_{10} , O_3 , SO_2 , NO_2 , CO dan $PM_{2.5}$.
2. Percobaan model membutuhkan dataset yang bersifat *time series* dan memiliki fitur yaitu waktu, *air quality index* (AQI), PM_{10} , O_3 , SO_2 , NO_2 , CO dan $PM_{2.5}$.
3. Prediksi dari model bersifat *continuous* untuk memprediksi kondisi udara di waktu yang akan mendatang berdasarkan pengenalan pola pada *dataset* yang digunakan. Hasil performa model dari pengembangan BiLSTM dengan

penambahan *convolutional* layer dilihat dari nilai *Mean Absolute Percentage Error* (MAPE).

1.6 Sistematika Penulisan

Sistematika penulisan pada penelitian ini memiliki lima BAB sebagai berikut.

1. BAB I PENDAHULUAN

Pada BAB I, berisi tentang pembahasan latar belakang penelitian beserta, rumusan masalah, tujuan, manfaat dan batasan masalah. Pembahasan BAB 1 merupakan dasar dari dilakukannya penambahan layer *convolutional* pada model BiLSTM. Kemudian, BAB 1 terdapat sistematika penulisan untuk penelitian yang dilakukan.

2. BAB II LANDASAN TEORI

Pembahasan BAB II mengenai teori yang mendukung penelitian. Teori pendukung beserta kajian penelitian terkait bertujuan untuk mengetahui dasar keilmuan dari penambahan layer *convolutional* pada model BiLSTM untuk memprediksi kualitas udara. Kajian terhadap teori dan penelitian terkait akan menjadi landasan konseptual penelitian yang dibahas di Bagian III.

3. BAB III METODOLOGI

Metodologi dipaparkan pada BAB III dengan pembahasan mengenai alur penelitian. Alur penelitian yang dilakukan yaitu data *collecting*, data *preprocessing*, *exploratory data analysis*, *Input data for modeling*, integrasi BiLSTM dengan *convolutional* layer dan evaluasi model.

4. BAB IV HASIL DAN PEMBAHASAN

BAB IV merupakan pembahasan mengenai hasil dari penelitian yang telah dilakukan sesuai alur penelitian pada BAB III. Hasil penelitian yang dibahas diantaranya yaitu hasil dari data *collecting*, *exploratory data analysis*, pemodelan BiLSTM diintegrasikan dengan layer *convolutional*, dan evaluasi model terhadap prediksi kualitas udara.

5. BAB V KESIMPULAN DAN SARAN

Pada BAB V, merangkum kesimpulan dari seluruh rangkaian penelitian yang telah dilakukan dan menyoroti kelebihan dan kelemahan penelitian. Kemudian, memberikan saran untuk peluang penelitian selanjutnya guna memperbaiki kekurangan yang ada dalam penelitian yang telah dilaksanakan.