

BAB I

PENDAHULUAN

Bab ini akan membahas mengenai pendahuluan yang mendasari dilakukannya kegiatan penelitian. Beberapa sub-bab yang dijelaskan dalam bagian ini yaitu latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah yang digunakan sebagai acuan selama melakukan penelitian.

1.1. Latar Belakang

Udara merupakan salah satu aspek krusial yang harus ada dalam kehidupan makhluk hidup. Udara bersih sangat penting karena berkontribusi besar pada kinerja kognitif dan kesehatan paru-paru, jantung, serta mental manusia[1]. Manfaat udara lebih dari sekadar kebutuhan dasar makhluk hidup untuk bernapas, melainkan juga memiliki peran yang signifikan terhadap proses ekologi lingkungan dengan membantu mempertahankan keseimbangan iklim. Kualitas udara mempunyai pengaruh yang besar terhadap kesehatan masyarakat, kelestarian lingkungan, dan pembangunan ekonomi. Namun demikian, peningkatan aktivitas industri dan urbanisasi telah menurunkan kualitas udara di berbagai wilayah, terutama di Indonesia.

Polusi udara merupakan masalah utama di Indonesia yang memiliki dampak negatif terhadap kualitas hidup masyarakat. Polusi udara terjadi ketika terdapat polutan berbahaya memasuki atmosfer, mengubah komposisi udara yang seharusnya bersih menjadi tercemar [2]. Polutan ini dapat berasal dari faktor alam, yaitu letusan gunung berapi, maupun dari aktivitas manusia, seperti emisi mobil, pembakaran bahan bakar fosil, serta limbah industri. *World Health Organization* (WHO) memperkirakan bahwa polusi udara bertanggung jawab atas sekitar 4,2 juta kematian dini setiap tahunnya, dengan materi partikulat (PM) sebagai kontributor utama penyakit pernapasan dan kardiovaskular [3]. Beberapa zat polutan udara yang sering dijumpai mencakup nitrogen oksida (NO₂), sulfur dioksida (SO₂), karbon monoksida (CO), senyawa organik yang mudah menguap (VOC), dan bahan partikulat [4].

Di lingkungan perkotaan, konsentrasi polutan udara tergolong bervariasi karena faktor-faktor seperti kondisi meteorologi, topografi, dan aktivitas manusia. Banyak kota besar di Indonesia memiliki tingkat polusi udara yang cukup tinggi, salah satunya adalah Kota Yogyakarta. Kualitas udara di Yogyakarta sering kali dipengaruhi oleh aktivitas industri, transportasi, dan faktor meteorologi. Menurut data dari Dinas Lingkungan Hidup Kota Yogyakarta, konsentrasi polutan seperti CO, SO₂, dan CO mengalami penurunan selama masa pandemi COVID-19 pada tahun 2020, dengan penurunan masing-masing sebesar 18%, 93%, dan 74% dibandingkan tahun 2019 [5]. Meskipun demikian, berdasarkan data *real-time* dari AQI.in, Indeks Kualitas Udara (AQI) di Yogyakarta mencapai kategori "Tidak Sehat" dengan konsentrasi PM_{2.5} dan CO yang melebihi ambang batas yang ditetapkan oleh *World Health Organization* (WHO).

Karbon monoksida (CO) adalah salah satu konsentrat berbahaya yang kebanyakan berasal dari pembakaran bahan bakar fosil. Gas CO dapat mengakibatkan beragam dampak serius bagi kesehatan manusia. Paparan CO dengan lebih dari 10% dapat menyebabkan gejala-gejala seperti sesak napas, pusing, kebingungan, sakit kepala, mual, kelelahan, nyeri dada, bahkan kehilangan kesadaran. Bahkan, apabila terjadi paparan jangka panjang maka dapat berujung pada gangguan kardiovaskular dan kematian akibat asfiksia. Terdapat penelitian yang menunjukkan bahwa anak-anak yang tinggal di daerah dengan tingkat paparan CO tinggi lebih rentan terhadap dampak negatif yang ditimbulkan terutama apabila dengan nilai hazard quotient jauh melebihi ambang batas aman yang ditetapkan oleh standar kesehatan lingkungan[6].

Oleh karena itu, sangat penting untuk membangun sistem yang mampu melakukan prediksi konsentrasi polusi udara secara tepat, sehingga memungkinkan adanya tindakan pencegahan oleh pihak berwenang. Seiring dengan pesatnya kemajuan *machine learning* dan pemanfaatan data, pengembangan terkait pendekatan yang lebih canggih untuk sistem *forecasting* mulai dilakukan. Algoritma seperti *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM) telah memperlihatkan kemampuannya dalam menangani data yang rumit dan tidak teratur. Jika dilihat dari penelitian yang dilakukan oleh Kim, pemanfaatan algoritma *Convolutional Neural Network* (CNN) dapat mendeteksi

pola-pola penting yang ada dalam data yang tidak dapat dideteksi menggunakan metode konvensional biasa [7]. Selain itu, penerapan algoritma *Long Short-Term Memory* (LSTM) juga memiliki hasil yang komprehensif dalam hal mengelola data *time-series*, terutama yang memiliki ketergantungan jangka panjang [8]. Sehingga, penggabungan kedua algoritma CNN-LSTM ini dapat memberikan hasil yang lebih baik pada sistem prediksi konsentrasi zat polutan udara. Model ini menghasilkan *output* yang lebih akurat dalam meramalkan konsentrasi zat polutan di udara dibandingkan dengan menggunakan pendekatan tunggal, seperti LSTM atau CNN saja [7].

Pada model prediksi ini, penambahan *lagged features* ke dalam model CNN-LSTM diharapkan dapat meningkatkan sensitivitas terhadap pola jangka panjang yang mungkin tidak terlihat jika hanya menggunakan model saja. Selain itu, penggabungan model prediksi CNN-LSTM dengan teknik *lagged features* dapat meningkatkan representasi temporal dari data, sehingga memungkinkan sistem untuk melihat relasi yang lebih kompleks antara konsentrasi polutan dan variabel-variabel lain seperti kelembapan udara, suhu, arah angin, dan kecepatan angin yang secara langsung terkait dengan tingkat polusi.

Keberadaan sistem prediksi ini nantinya diharapkan dapat membantu pemerintah dalam membuat keputusan terutama terkait dengan kebijakan emisi serta pengelolaan limbah-limbah berbahaya. Hasil yang akurat terkait dengan jumlah zat polutan yang ada di udara memungkinkan lembaga berwenang serta pemerintah untuk membuat keputusan yang diperlukan dalam menjaga kesehatan masyarakat dan lingkungan.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang sebelumnya telah dijabarkan, berikut merupakan rumusan masalah dalam penelitian ini:

1. Bagaimana implementasi teknik data *preprocessing lagged features* dalam model prediksi konsentrasi zat polutan udara pada data *time series*?
2. Bagaimana implementasi dari algoritma CNN-LSTM dalam melakukan prediksi konsentrasi zat polutan udara pada data *time series*?

3. Bagaimana penerapan dari algoritma CNN-LSTM dengan *lagged features* dalam melakukan prediksi konsentrasi zat polutan udara dapat ditingkatkan?

1.3. Tujuan Penelitian

Adapun tujuan penelitian berdasarkan rumusan masalah di atas, antara lain sebagai berikut:

1. Mengimplementasikan dan mengembangkan penerapan teknik data *preprocessing lagged features* untuk melakukan prediksi terhadap konsentrasi zat polutan udara pada data *time series*.
2. Mengimplementasikan dan mengembangkan penerapan algoritma CNN-LSTM untuk melakukan prediksi terhadap konsentrasi zat polutan udara pada data *time series*.
3. Mengoptimalkan implementasi dari algoritma CNN-LSTM dengan *lagged features* untuk memprediksi konsentrasi zat polutan udara dalam upaya untuk memberikan hasil akurasi yang lebih baik.

1.4. Manfaat Penelitian

Penelitian yang dilakukan mempunyai manfaat untuk masyarakat, baik masyarakat akademik maupun non akademik. Adapun beberapa manfaat dari penelitian ini, yaitu:

A. Manfaat bagi Masyarakat Akademik

1. Menambah wawasan dan referensi di bidang prediksi konsentrasi zat polutan udara menggunakan algoritma CNN-LSTM.
2. Mengembangkan ilmu pengetahuan di bidang *machine learning* dan analisis data *time-series*.
3. Sebagai acuan untuk penelitian lanjutan di bidang terkait.

B. Manfaat bagi Masyarakat Non-Akademik

1. Membantu instansi terkait dalam merencanakan kebijakan terkait dengan polusi udara.
2. Menjadi sarana pemantauan baik bagi masyarakat maupun pemerintah setempat.

3. Membantu adanya perencanaan mitigasi jangka panjang terkait problematika yang menyangkut kualitas udara.
4. Mendukung upaya pelestarian lingkungan melalui pemantauan konsentrasi zat polutan udara yang lebih optimal.

1.5. Batasan Masalah

Dalam penelitian ini, terdapat beberapa poin utama terkait dengan bahasan masalah yang diterapkan. Batasan-batasan masalah ini, meliputi:

1. Dataset yang digunakan adalah Data Konsentrasi Rata - Rata Harian AQMS Tahun 2020-2023 DI Yogyakarta yang bersifat *open-source*.
2. Terdapat dataset suhu udara, kelembapan udara, curah hujan, serta kecepatan & arah angin harian yang didapat dari Stasiun Klimatologi DI Yogyakarta yang dapat langsung diakses melalui *website* Data Online BMKG.
3. Model akan memberikan hasil prediksi berupa konsentrasi zat polutan udara karbon monoksida (CO) yang didapatkan dari pengolahan 7 hari data sebelumnya.