

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perubahan iklim global semakin meningkat akibat emisi gas rumah kaca, salah satunya gas metana (CH_4), yang memiliki efek pemanasan 28 kali lebih kuat dibandingkan karbon dioksida (CO_2) dalam kurun waktu 100 tahun (Darandary dkk., 2023). Salah satu sumber utama emisi metana adalah sampah yang tidak dikelola dengan baik, terutama dari Tempat Pembuangan Akhir (TPA) yang mengalami dekomposisi anaerobik (Vaiškūnaitė & Zagorskis, 2024).

Beberapa penelitian telah dilakukan dalam bidang prediksi jumlah sampah dan estimasi emisi gas metana. (Loni Manikari dkk., 2024) mengimplementasikan LSTM untuk prediksi volume sampah dengan MSE 0,02207, namun belum mengonversi hasil prediksi ke estimasi emisi metana dan hanya menggunakan *optimizer* Adam tanpa eksplorasi *optimizer* lain. (Sinaga dkk., 2023) melakukan estimasi produksi gas metana di TPA Talang Gulo menggunakan pendekatan kuantitatif deskriptif dengan faktor konversi IPCC, namun tidak menggunakan model prediktif berbasis *machine learning*. (Huang dkk., 2020) membandingkan *optimizer* AdaGrad dan RMSProp pada prediksi limbah konstruksi dengan LSTM, namun tidak melakukan *tuning hyperparameter* secara sistematis menggunakan *Grid Search*. (Fadhel dkk., 2024) menggunakan e-LSTM dengan *optimizer* RAdam dan mencapai $R^2 = 0.90$, namun tidak melakukan konversi ke estimasi emisi metana dan terbatas pada satu lokasi *landfill*.

Penelitian ini mengisi celah studi sebelumnya melalui keunggulan metodologis dan cakupan data. Dataset yang digunakan berskala nasional, berisi 1.699 entri dari 317 kabupaten/kota, melampaui penelitian terdahulu yang umumnya terbatas pada skala regional. Penelitian ini tidak hanya memprediksi timbunan sampah, tetapi juga mengonversi hasil prediksi menjadi estimasi emisi metana menggunakan faktor IPCC, pendekatan yang belum diterapkan pada studi (Loni Manikari dkk., 2024) maupun (Fadhel dkk., 2024). Proses *tuning* dilakukan secara sistematis memakai *Grid Search* dengan 243 kombinasi *hyperparameter*, berbeda dari (Huang dkk., 2020) yang hanya membandingkan dua *optimizer* dan (Loni Manikari dkk., 2024) yang menggunakan satu konfigurasi. Hasil evaluasi menunjukkan RMSProp sebagai *optimizer* terbaik dengan nilai R^2 0.9023, lebih tinggi dibandingkan Adam dan RAdam pada konfigurasi serupa. Tahap *preprocessing* mencakup normalisasi *Min-Max*, deteksi *outlier*, penanganan *missing values*, dan transformasi *sliding window* untuk memastikan kualitas *input data*.

Penelitian ini memiliki relevansi tinggi karena mendukung pencapaian SDGs, khususnya Tujuan 11 dan 13, serta memberikan kontribusi metodologis melalui integrasi prediksi *time series* dan estimasi dampak lingkungan dalam satu kerangka terpadu. Penelitian juga menghasilkan model yang dapat dimanfaatkan untuk perencanaan strategis pengelolaan sampah dan mitigasi emisi gas rumah kaca di tingkat nasional. Penelitian ini menunjukkan adanya celah signifikan dalam integrasi prediksi timbunan sampah dan estimasi emisi metana yang dapat diisi melalui pendekatan *deep learning* yang lebih sistematis dan komprehensif.

Prediksi emisi gas metana dari sampah menjadi langkah penting untuk membantu pemerintah menyusun strategi pengurangan emisi gas rumah kaca dan memitigasi perubahan iklim. Relevansi topik ini sejalan dengan pencapaian *Sustainable Development Goals* (SDGs), khususnya dengan Tujuan 11 (Kota dan Pemukiman yang Berkelanjutan) serta Tujuan 13 (Penanganan Perubahan Iklim). SDGs merupakan serangkaian 17 tujuan global yang ditetapkan oleh Perserikatan Bangsa-Bangsa (PBB) pada tahun 2015 sebagai bagian dari Agenda 2030 untuk Pembangunan Berkelanjutan (Senger, 2017).

Tantangan prediksi data *time series* dengan pola kompleks dan dependensi temporal membutuhkan model yang mampu menangkap hubungan jangka panjang dalam data historis. *Long Short-Term Memory* (LSTM) merupakan arsitektur *deep learning* yang dirancang khusus untuk mengatasi permasalahan *long-term dependency* dalam data deret waktu melalui mekanisme *gating* yang memungkinkan model mempertahankan informasi penting dalam jangka waktu yang panjang (Fukai dkk., 2024).

Beberapa varian dari LSTM, seperti *Bidirectional LSTM* (BiLSTM) dan *Enhanced LSTM* (e-LSTM), telah dikembangkan untuk meningkatkan akurasi prediksi dan menangkap pola yang lebih kompleks dalam data *time series* (Pokkuluri & Khang, 2025). Model-model tersebut menunjukkan performa yang menjanjikan, namun tantangan seperti *overfitting* dan pemilihan *hyperparameter* yang optimal tetap menjadi hambatan dalam implementasi praktisnya (Diukarev & Starukhin, 2024; Hwang dkk., 2024).

Pemilihan *optimizer* memengaruhi kinerja dan risiko *overfitting*. Studi sebelumnya menunjukkan bahwa Adam dan RAdam dapat mempercepat konvergensi, namun keduanya cenderung meningkatkan risiko *overfitting* karena terlalu cepat mencapai minimum lokal (Wilson dkk., 2017; M. R. Zhang dkk., 2019). RMSProp menawarkan keseimbangan yang lebih baik antara kecepatan dan kemampuan generalisasi karena menyesuaikan *learning rate* secara dinamis berdasarkan rata-rata kuadrat gradien sebelumnya (Mukkamala & Hein, 2017). Mekanisme ini mencegah osilasi pada area curam dan mempercepat konvergensi pada area datar, sehingga pelatihan menjadi lebih stabil dan efisien (Q. Zhang dkk., 2020). Konteks prediksi timbulan sampah yang memiliki pola musiman dan tren jangka panjang menuntut mekanisme adaptif, sehingga karakteristik RMSProp membuatnya lebih unggul dan mampu menghasilkan model yang akurat serta memiliki generalisasi baik pada data uji.

Pentingnya *tuning hyperparameter* secara sistematis telah dibuktikan dalam berbagai studi *time series*. (Rong dkk., 2022) menunjukkan bahwa optimasi struktur dan *hyperparameter* LSTM dapat meningkatkan akurasi prediksi konsumsi listrik, sementara (Huang dkk., 2020) menemukan perbedaan performa signifikan antara *optimizer* pada prediksi limbah konstruksi. Banyak penelitian masih menggunakan pendekatan manual dalam memilih konfigurasi model. *Grid Search* menawarkan eksplorasi yang lebih menyeluruh dan objektif terhadap ratusan kombinasi parameter, sehingga mengurangi bias subjektif (Anggreani, 2024; Khullar, 2024). Penggunaan *Grid Search* memastikan model tidak hanya akurat pada data pelatihan, tetapi juga memiliki kemampuan generalisasi yang kuat pada data uji

dalam konteks prediksi timbunan sampah yang kompleks dan bervariasi antar wilayah.

Penelitian ini memperkuat pendekatan prediksi timbunan sampah melalui beberapa inovasi. Prediksi jumlah sampah diintegrasikan dengan estimasi emisi metana menggunakan faktor IPCC sebesar 50 kg CH₄ per ton sampah berdasarkan (Rarastry, 2016), suatu langkah yang belum diterapkan pada penelitian sejenis. Strategi pengembangan model mencakup pemilihan *optimizer* RMSProp, penerapan *dropout*, serta *tuning hyperparameter* melalui *Grid Search*. Tahap *preprocessing* dilakukan secara cermat melalui normalisasi, deteksi *outlier*, penanganan *missing values*, dan *sliding window* untuk meningkatkan kualitas *input*. Dataset nasional yang digunakan mencakup 1.699 entri dari 317 kabupaten/kota periode 2018–2024, memberikan cakupan lebih luas dibandingkan penelitian terdahulu. Pendekatan yang sistematis serta integrasi dengan estimasi emisi metana memberikan kontribusi bagi pengelolaan sampah berkelanjutan dan perumusan kebijakan mitigasi perubahan iklim berbasis data.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, penelitian ini berusaha menjawab pertanyaan-pertanyaan berikut:

1. Bagaimana mengimplementasikan model LSTM untuk memprediksi jumlah sampah menggunakan data historis harian dan tahunan melalui penerapan teknik regularisasi *dropout*, pemilihan *optimizer* RMSprop, serta *tuning hyperparameter* menggunakan *Grid Search* guna meningkatkan akurasi dan mencegah *overfitting*?

2. Bagaimana menghitung estimasi emisi gas metana berdasarkan hasil prediksi jumlah sampah menggunakan faktor konversi IPCC 1 ton sampah = 50 kg gas metana (CH_4)?
3. Bagaimana mengukur performa model LSTM dalam memprediksi jumlah sampah dan estimasi emisi gas metana menggunakan metrik evaluasi *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *R-Squared* (R^2)?

1.3 Tujuan Penelitian

Sesuai dengan rumusan masalah di atas, tujuan dari penelitian ini adalah:

1. Mengimplementasikan model LSTM untuk memprediksi jumlah sampah menggunakan data historis harian dan tahunan melalui penerapan teknik regularisasi *dropout*, pemilihan *optimizer* RMSprop, serta *tuning hyperparameter* menggunakan *Grid Search* guna meningkatkan akurasi dan mencegah *overfitting*.
2. Menghitung estimasi emisi gas metana berdasarkan hasil prediksi jumlah sampah menggunakan faktor konversi IPCC 1 ton sampah = 50 kg gas metana (CH_4).
3. Mengukur performa model LSTM dalam memprediksi jumlah sampah dan estimasi emisi gas metana menggunakan metrik evaluasi *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *R-Squared* (R^2).

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah referensi dalam bidang *data science* dan *deep learning*, khususnya pada prediksi *time series* menggunakan model LSTM.
2. Memberikan wawasan baru mengenai pemanfaatan prediksi data dalam pengelolaan sampah dan mitigasi dampak lingkungan.
3. Menjadi acuan bagi pemerintah dan pengelola sampah dalam memprediksi jumlah sampah dan emisi gas metana di masa depan.
4. Mendukung perencanaan strategi pengelolaan sampah yang lebih efektif berdasarkan hasil prediksi.
5. Menawarkan pendekatan *machine learning* sebagai alat bantu *monitoring* dan prediksi jumlah sampah serta estimasi emisi gas metana.
6. Memberikan model prediksi sebagai referensi awal bagi pengembangan sistem peringatan dini emisi metana dalam penelitian lanjutan di masa depan.

1.5 Batasan Penelitian

Penelitian lebih terfokus, terdapat beberapa batasan yang diterapkan:

1. Data yang digunakan berupa dataset jumlah sampah harian dan tahunan dari Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) dalam format CSV, yang mencakup data historis dalam rentang waktu tertentu.
2. Emisi gas metana dihitung berdasarkan formula sederhana faktor konversi IPCC (1 ton sampah = 50 kg gas metana) tanpa mempertimbangkan faktor lain seperti jenis sampah, laju dekomposisi, atau kondisi lingkungan.
3. Model dievaluasi menggunakan metrik MAE, MSE, RMSE, dan R^2 untuk mengukur performa model.