

BAB I

PENDAHULUAN

1.1. Latar Belakang

Udara merupakan salah satu komponen yang paling penting untuk kelangsungan makhluk hidup terutama manusia. Menurut Hesam (2005) dalam Naddafi dkk (2006), kebutuhan udara bagi manusia lebih utama daripada kebutuhan terhadap makanan dan air. Rata-rata kebutuhan udara orang dewasa adalah 15 kg/hari, sedangkan kebutuhan makanan dan air masing-masing sebesar 1.5 kg/hari dan 2.5 kg/hari. Manusia dapat terus hidup tanpa makanan selama lima minggu dan tanpa air selama lima hari, tetapi tidak lebih dari beberapa menit tanpa udara. Namun, kondisi udara di perkotaan saat ini sudah berubah dari batas normal. Hal ini disebabkan oleh polusi udara seperti dari sisa hasil pembakaran mesin kendaraan bermotor atau emisi gas buang, kegiatan industri atau kegiatan lain yang mengakibatkan udara tercemar.

Pencemaran udara dikenal sebagai masalah lingkungan yang terasosiasi dengan wilayah perkotaan di seluruh dunia. Pencemaran udara juga menjadi salah satu indikator kualitas lingkungan yang berdampak pada kesehatan masyarakat dan dapat mempengaruhi kualitas udara di wilayah tersebut. Menurut Setiawan (1992), terjadinya pencemaran udara luar ruangan dapat menyebabkan berbagai penyakit saluran pernapasan seperti batuk, bronkhitis, asma, pneumonia, dan bahkan kanker paru. Oleh karena itu, berbagai program pemantauan telah

dilakukan untuk menentukan kualitas udara dengan menggeneralisasikan sejumlah data dari setiap konsentrasi polutan (Wu dan Kuo, 2013).

Menurut Laporan *State of Global Air* tahun 2018 yang diterbitkan oleh *Health Effects Institute* di Amerika dan *Institute for Health Metrics and Evaluation's Global Burden of Disease Project*, polusi udara merupakan penyebab terbesar ke-6 yang dapat memperpendek usia manusia atau kematian dini. Jumlah korban dapat melebihi kematian yang disebabkan faktor risiko lain seperti kecanduan alkohol, aktivitas fisik, konsumsi Natrium yang tinggi atau kolesterol. Namun, tingkat kematian dini tersebut tidak sama rata di semua negara. Indonesia merupakan salah satu dari negara asia yang terpengaruh oleh polusi udara. Kelima negara dengan tingkat kematian paling tinggi terdapat di Asia yaitu China, India, Pakistan, Indonesia dan Bangladesh. Organisasi Kesehatan Dunia (WHO) mengatakan bahwa pada tahun 2016 polusi udara baik di perkotaan maupun pedesaan menyebabkan 4,2 juta kematian di seluruh dunia dan 3,8 juta kematian lain akibat penyakit tidak menular (jantung, stroke dan kanker) yang disebabkan oleh polusi udara rumah tangga. Dinyatakan bahwa 91% populasi di dunia tidak menghirup udara bersih dan lebih dari setengah populasi di daerah perkotaan terpapar oleh polusi udara luar ruangan setidaknya 2,5 kali di atas standar keamanan yang ditetapkan oleh WHO (WHO, 2018).

Sektor transportasi sebagai tulang punggung aktivitas masyarakat berkontribusi besar pada pencemaran udara karena mayoritas kendaraan bermotor masih menggunakan bahan bakar fosil. Menurut Badan Pusat Statistik, jumlah kendaraan bermotor di Indonesia yang mencakup mobil penumpang, mobil bis,

mobil barang dan sepeda motor dalam kurun waktu 10 tahun terakhir (2007-2017) mengalami pertumbuhan yang sangat tinggi yakni sebesar 153%. Sepeda motor mendominasi dengan 82% di tahun 2017. Seiring meningkatnya jumlah kendaraan bermotor yang beredar di masyarakat menyebabkan meningkatnya polutan udara yang berupa emisi gas buang. Emisi gas buang merupakan gas yang dibuang oleh kendaraan bermotor dari sisa hasil proses pembakaran bahan bakar di dalam mesin. Emisi dari kendaraan bermotor yang terus-menerus terpapar di udara menghasilkan berbagai macam polutan, salah satunya adalah *particulate matter* 10 (PM10) sebagai akibat proses pembakaran mesin yang tidak sempurna serta partikel lepas (Syahrani, 2006).

Dinas Lingkungan Hidup DKI Jakarta mencatat konsentrasi PM10 di udara Jakarta selama lima tahun terakhir mencapai 60,8 mikrogram per meter kubik. Angka tersebut jauh di atas jumlah standar konsentrasi udara menurut Badan Kesehatan Dunia (WHO) yakni 20 mikrogram per meter kubik dalam jangka waktu satu tahun (Tresnasari, 2018). Polusi udara karena emisi gas buang ini menjadi salah satu masalah yang muncul seiring berkembangnya sebuah kota menuju kota maju. Sebuah kota dianggap maju jika mampu memanfaatkan sumber daya yang ada secara cerdas.

Smart city menjadi label yang melekat di sebuah kota untuk menjawab tantangan dari berbagai macam masalah yang muncul di masyarakat untuk dipecahkan. Istilah *smart city* telah menjadi tren tersendiri dalam pelayanan di pemerintah daerah, baik di tingkat kota maupun kabupaten. *Smart city* berarti bahwa sebuah kota mampu memanfaatkan modal yang dimiliki, seperti sumber

daya manusia, kekuatan hubungan sosial, dan infrastruktur telekomunikasi modern untuk mewujudkan pertumbuhan ekonomi berkelanjutan dan kualitas kehidupan yang tinggi, dengan manajemen sumber daya yang bijaksana melalui pemerintahan berbasis partisipasi masyarakat (Caragliu *et al.*, 2009). Kota pintar tercipta dengan tujuan untuk meningkatkan kenyamanan penghuninya melalui fasilitas mobilitas, efisiensi dan hemat energi, perbaikan kualitas air dan udara, identifikasi masalah dan pemberian solusi cepat, tanggap dalam pemulihan setelah bencana, dan integrasi data untuk pengambilan keputusan melalui pemanfaatan *cloud computing* dalam mengolah *big data* (Benevolo, 2016; Bhati, 2017; Suakanto, 2013; Suri, 2018; Nuaimi, 2015).

Smart environment merupakan aspek *smart city* yang bertujuan untuk melindungi dan memelihara lingkungan kota, baik keamanan maupun alam yang fokus pada pembahasan kemajuan teknologi serta pemanfatannya. Lingkungan pintar berarti lingkungan yang mampu menghadirkan kenyamanan bagi masyarakat. Salah satu cara mengembangkan lingkungan pintar adalah penggunaan teknologi canggih untuk memproses data sehingga melahirkan informasi dan menganalisisnya (Kumar, 2020). Jakarta merupakan salah satu kota di Indonesia yang telah menerapkan *smart city*, terutama di aspek *smart environment*. Melalui Dinas Lingkungan Hidup, Pemerintah Kota Jakarta telah melakukan pemantauan kualitas udara ambien pada beberapa parameter udara yaitu CO, PM10, O3, NOx, NO2 dan SO2. Data hasil pemantauan tersebut disimpan dalam bentuk file yang volume-nya tiap hari bertambah besar. Dengan volume data dan varian yang besar ini memungkinkan untuk diolah sehingga

memberikan pengetahuan baru untuk menangani permasalahan lingkungan. Misalnya dengan mengimplementasikan teknologi *artificial intelligence*.

Saat ini, *Artificial Intelligence* (AI) makin masif diaplikasikan dalam kehidupan masyarakat untuk membantu penyelesaian masalah. Salah satu konsepsi AI yang memanfaatkan data besar untuk menyelesaikan masalah adalah *forecasting*. Peramalan (*forecasting*) adalah suatu teknik analisis perhitungan untuk memperkirakan kejadian di masa yang akan datang dengan menggunakan pengalaman di masa lampau (Sumayang, 2013). *Forecasting* merupakan bagian terpenting bagi setiap perusahaan, organisasi ataupun pemerintahan dalam setiap pengambilan keputusan manajemen. Peramalan bisa menjadi dasar bagi para pembuat kebijakan untuk membuat perencanaan jangka pendek, menengah maupun jangka panjang bagi organisasi komersial maupun tidak. Salah satu algoritma yang dipakai dalam peramalan yang memanfaatkan *Artificial Intelligence* dan berbasis data runtun waktu adalah *Long Short Term Memory* (LSTM) yang merupakan jenis dari *Recurrent Neural Network* (RNN).

Recurrent Neural Network (RNN) atau jaringan saraf berulang merupakan bentuk pengembangan dari *Artificial Neural Network* (ANN). Pengembangan tersebut dilakukan karena ANN tidak mampu melakukan proses *training* secara cepat pada data yang banyak. RNN sendiri dirancang untuk memproses data bersambung atau berurutan, serta biasa digunakan untuk menyelesaikan masalah dengan data *time series*. Ciri khas RNN adalah data yang dijadikan masukan adalah data sekuensial. Karena data diproses melalui banyak lapis (layer) sehingga RNN masuk dalam kategori *deep learning*. Seiring dengan berkembangnya

teknologi *hardware* untuk membantu proses *deep learning*, RNN telah mengalami kemajuan yang pesat dan telah diterapkan ke dalam berbagai bidang. Seiring berkembangnya implementasi RNN, dikembangkan pula metode *Long Short Term Memory* (LSTM).

Long Short Term Memory (LSTM) merupakan bentuk pengembangan RNN yang dimaksudkan untuk menghindari masalah ketergantungan jangka panjang dan sesuai untuk memproses serta memprediksi deret waktu. Berbeda dengan konsep RNN, LSTM memiliki *memory cell* dan *input gates* untuk menyimpan informasi memori yang telah lama tersimpan. LSTM banyak dipilih untuk prediksi berbasis waktu atau *time series* karena dikenal lebih unggul dan handal dalam melakukan prediksi dalam waktu lama dibanding algoritma lain (Zahara, Sugianto, & Ilmiddafiq, 2019).

Disamping LSTM sebagai alat yang ampuh dalam *time series* dan pengenalan pola terdapat beberapa kekurangan dalam penggunaan LSTM *network*. Pertama, *neural network models* termasuk LSTM kurang mampu dalam menjelaskan keputusan akhir yang diperoleh model. Model *neural network* ini mempunyai proses komputasi yang sangat kompleks, yang dapat mencapai solusi yang tepat dalam menyelesaikan masalah, tetapi tidak dapat menjelaskan secara spesifik hasil dari prediksinya. Oleh karena itu, Shin dan Lee (2002) menggunakan *hybrid* antara *Genetic Algorithm* (GA) dengan *Artificial Neural Network* (ANN) dalam memprediksi kebangkrutan suatu perusahaan. *Genetic Algorithm* (GA) merupakan salah satu algoritma optimasi yang diciptakan untuk meniru beberapa proses yang diamati dalam evolusi alam dan genetika alam

yang mampu mengeksplorasi pencarian dalam jumlah besar secara efektif, sehingga memperoleh hasil terbaik (Guo *et al*, 2010). Kedua, seperti *neural network models* yang lain LSTM mempunyai banyak parameter yang harus dimodifikasi seperti jumlah *layers*, *neuorns per layer* dan jumlah *time lags* (Chung, 2018).

Penelitian terdahulu menggunakan LSTM diantaranya Zheng dkk (2017), dalam penelitian ini model LSTM dibandingkan dengan beberapa model peramalan lain yaitu dengan SARIMA, NARX, SVR, dan NNETAR. Dihadirkan bahwa model LSTM mampu mengungguli metode peramalan lain dengan menghasilkan nilai RMSE dan MAPE yang lebih rendah. Nilai RMSE yang dihasilkan yaitu sebesar 0,0702, sementara nilai MAPE yang dihasilkan sebesar 0,0535%. Aulia (2020) tentang prediksi harga Ethereum dengan menggunakan metode LSTM menunjukkan bahwa model prediksi terbaik diperoleh dengan jumlah neuron 50 dan *max epoch* 500 serta nilai MAPE sebesar 1.69 %. Terdapat juga penelitian oleh Sugiarto (2017) mengenai optimasi *Artificial Neural Network* menggunakan *Genetic Algorithm* pada prediksi *approval credit card* dengan menggunakan algoritma *neural network* mendapatkan hasil peningkatan dari 85.42% menjadi 87.82% hasil optimasi dengan *Genetic Algorithm*. Chung dkk (2018) melakukan penelitian tentang prediksi mengenai harga saham KOSPI dengan menggunakan *Genetic Algorithm-Long Short Term Memory* (GA-LSTM) yang menghasilkan bahwa *time window size* yang diperoleh dengan optimasi GA sebesar 10 dan menggunakan dua *hidden layer* dengan jumlah *node* 15 dan 7.

Berdasarkan uraian diatas, dalam penelitian ini akan dilakukan *hybrid* antara GA dengan LSTM dalam memprediksi kualitas udara kota Jakarta yang selanjutnya akan dibandingkan antara metode LSTM dengan GA-LSTM. Melalui penelitian ini, diharapkan dapat diperoleh model terbaik yang dapat memberikan hasil prediksi yang akurat mengenai kualitas udara di Kota Jakarta dan dapat dijadikan sebagai salah satu upaya untuk meningkatkan sistem pemantau kualitas udara yang ada di Kota Jakarta. Dari penelitian ini diharapkan para pemangku kepentingan khususnya di bidang pengendalian polusi lingkungan, dapat menggunakan metode ini sebagai acuan pengambilan keputusan dalam upaya mengendalikan polusi udara untuk tahun-tahun yang akan datang.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang di atas maka dapat dirumuskan suatu permasalahan yakni:

1. Bagaimana karakteristik dari data kadar polutan udara di Kota Jakarta?
2. Bagaimana pemodelan data polutan udara kadar polutan udara di Kota Jakarta menggunakan *Long Short-Term Memory* (LSTM) dan *Genetic Algorithm Long Short-Term Memory* (GA-LSTM)?
3. Bagaimana perbandingan tingkat akurasi dari pemodelan data kadar polutan udara di Kota Jakarta menggunakan *Long Short-Term Memory* (LSTM) dan *Genetic Algorithm Long Short-Term Memory* (GA-LSTM)?

1.3. Tujuan Penelitian

Berdasarkan permasalahan di atas, maka didapatkan tujuan dari penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik data kadar polutan udara di Kota Jakarta.
2. Mendapatkan model data kadar polutan udara di Kota Jakarta menggunakan *Long Short-Term Memory* (LSTM) dan *Genetic Algorithm Long Short-Term Memory* (GA-LSTM).
3. Membandingkan akurasi dari pemodelan data kadar polutan udara di Kota Jakarta menggunakan *Long Short-Term Memory* (LSTM) dan *Genetic Algorithm Long Short-Term Memory* (GA-LSTM).

1.4. Manfaat Penelitian

Adapun manfaat yang bisa dipetik dari penelitian ini diantaranya adalah:

1. Manfaat Teoritis

Membantu perkembangan ilmu pengetahuan mengenai *Artificial Neural Network* (ANN) sehingga dapat digunakan sebagai bahan bacaan dan referensi bagi pembaca dalam melakukan penelitian terutama yang berhubungan dengan metode *Long Short Term Memory* dan *Genetic Algorithm*.

2. Manfaat Praktis

- a. Bagi Peneliti, peneliti mampu menerapkan metode yang sesuai dalam materi yang telah dipelajari serta peneliti mempunyai pengetahuan dan wawasan mengenai *Artificial Neural Network* (ANN) untuk peramalan

menggunakan metode *Long Short Term Memory* dan *Long Short Term Memory* dengan *Genetic Algorithm*.

- b Bagi pihak terkait diantaranya pemerintah dan masyarakat dapat mengetahui kadar polutan udara untuk beberapa periode dan sebagai salah satu bahan pertimbangan pemerintah dalam mengambil kebijakan khususnya masalah yang berkaitan dengan pencemaran udara khususnya di Jakarta.

1.5. Batasan Masalah

Penelitian ini difokuskan pada data per jam kadar polutan udara di Kota Jakarta pada tanggal 1 Januari hingga 30 November 2020. Metode yang digunakan dalam penelitian ini adalah *Long Short-Term Memory* (LSTM) dan *Genetic Algorithm-Long Short-Term Memory* (GA-LSTM). GA digunakan untuk mencari *window size* dan *number of units* yang optimal pada arsitektur LSTM. Hasil peramalan akan dibandingkan menggunakan nilai MAPE.