

BAB I

PENDAHULUAN

A. Latar Belakang

Udara merupakan kebutuhan dasar bagi keberlangsungan hidup dan perkembangan semua kehidupan di Bumi (Bai, et al. 2018), mulai dari manusia, tumbuhan, binatang, bahkan nilai estetika visual dapat dipengaruhi oleh kualitas udara. Udara yang tercemar dapat mempengaruhi kehidupan suatu populasi bahkan membahayakan kelangsungan hidup umat manusia. Pencemaran udara di daerah perkotaan yang pada umumnya disebabkan oleh penggunaan kendaraan bermotor secara intens khususnya mobil pribadi dan kendaraan barang berat merupakan masalah prioritas bagi perencanaan transportasi dan otoritas publik (Catalano, et al. 2016). Peningkatan penggunaan batu bara selama revolusi industri oleh pabrik dan rumah tangga juga menyebabkan kabut asap sehingga kesejahteraan populasi secara fisik, mental, maupun sosial menurun sedangkan angka rata-rata kematian meningkat secara signifikan.

DKI Jakarta yang dinobatkan sebagai wilayah metropolitan terbesar di Asia Tenggara atau urutan kedua di dunia ini memiliki luas sekitar 6392 km² dengan jumlah penduduk sebanyak 10,7 juta jiwa pada tahun 2020 (IQAir n.d.). Dengan jumlah populasi yang besar, jalanan dipenuhi dengan sepeda motor, mobil, bahkan truk yang dapat menghasilkan tingkat emisi gas yang tinggi. Polutan seperti nitrogen dioksida (NO₂) dan sulfur dioksida (SO₂) menjadi senyawa utama yang terlibat dalam penggunaan kendaraan. Selain kendaraan, bahan bakar fosil lain yang dibakar untuk menghasilkan energi juga dapat merusak kualitas udara. Sejumlah polutan seperti karbon monoksida (CO), karbon hitam, ozon (O₃), dan senyawa organik volatil (VOC'S) dilepaskan ke atmosfer sehingga dapat meningkatkan persentase emisi yang besar.

Tingkat polusi yang tinggi dapat memberikan dampak terhadap kesehatan dengan beberapa gejala seperti iritasi pada mata, kulit, dan hidung

serta penyakit pernapasan seperti bronkitis ataupun asma. Dampak lain yang mungkin terjadi adalah kerusakan pada jantung dan sistem peredaran darah karena ukuran polutan yang sangat kecil dapat masuk ke dalam aliran darah dan menyebabkan kerusakan pada pembuluh darah, meningkatkan kemungkinan terjadinya serangan jantung, serta meningkatkan risiko penyakit jantung ataupun gangguan jantung lainnya. Dampak-dampak tersebut merupakan sebagian efek yang dapat diderita seseorang ketika terpapar udara yang kualitasnya buruk dalam jangka waktu yang lebih lama.

Keputusan Menteri Negara Lingkungan Hidup Nomor 45 Tahun 1997 tentang Indeks Standar Pencemar Udara ditetapkan untuk memberikan keseragaman informasi yang dipublikasikan kepada masyarakat pada lokasi dan waktu tertentu serta sebagai bahan pertimbangan dalam melakukan upaya pengendalian terhadap pencemaran udara. Indeks Standar Pencemar Udara (ISPU) merupakan angka yang tidak mempunyai satuan dan menggambarkan kualitas udara ambians berdasarkan dampak terhadap nilai estetika, kesehatan manusia, dan makhluk hidup lainnya. Indeks Standar Pencemar Udara (ISPU) dihitung berdasarkan beberapa nilai parameter, diantaranya yaitu partikular (PM10), karbon monoksida (CO), sulfur dioksida (SO₂), nitrogen dioksida (NO₂), dan ozon (O₃). Hasil perhitungan ISPU akan menghasilkan nilai pada rentang tertentu yang sudah dikategorikan sesuai dengan lampiran Kepala Bapedal No. 107 tahun 1997, yaitu sebagai kualitas udara yang baik, sedang, tidak sehat, sangat tidak sehat, atau berbahaya.

Data Indeks Standar Pencemar Udara (ISPU) merupakan data runtun waktu (*time series*), yaitu sekumpulan data yang diambil secara bersamaan dalam interval waktu tertentu. Deret waktu *univariate* adalah data dengan variabel terikat waktu yang jumlahnya tunggal, sedangkan deret waktu *multivariate* merupakan data yang memiliki beberapa variabel tergantung waktu yang digunakan. Model *Autoregressive* (AR) merupakan model regresi yang menghubungkan nilai amatan dengan nilai-nilai sebelumnya pada rentang waktu tertentu, sedangkan model *Nonlinear Autoregressive* (NAR) merupakan model AR yang menggunakan fungsi *nonlinear* sebagai fungsi autoregresinya (Dietz 2010). Fungsi *nonlinear* dapat didekati oleh model

yang bersifat *nonlinear*, salah satu model yang bersifat *nonlinear* adalah *Neural Network*.

Salah satu arsitektur *neural network* yang paling sering digunakan adalah *multilayer perceptron* yang berisi satu atau lebih lapisan (*layer*) tersembunyi serta membangun model neuron dalam jaringan agar dapat memiliki fungsi aktivasi *nonlinear*. *Recurrent Neural Network* (RNN) menggabungkan *multilayer perceptron* statis dan mengeksplorasi kemampuan pemetaan *nonlinear* dari *perceptron* tersebut dengan mengumpulkan kembali nilai melalui elemen *time delay* ke masing-masing neuron dalam jaringan. Model *Recurrent Nonlinear Autoregressive with Exogenous Inputs* (NARX) *Neural Networks* merupakan salah satu model dari RNN yang elemen *time-delay* nya berupa satu set (dua garis) *tapped-delay*, satu *tap* terdiri dari nilai *input* dan *tap* lain pada nilai *output* (Siegelmann, Horne and Giles 1997). *Tapped-delay* pada NARX menjadi implementasi dari penyisipan *delay space* dan dapat menjadi dasar model statistik dari *Autoregressive* (AR) tradisional (Lin, et al. 1997). Pada penelitian sebelumnya, (Rege, et al. 2018) menggunakan pendekatan model ini untuk menyediakan kerangka kerja dalam memprediksi bagaimana musuh bergerak seiring kemajuan serangan siber dengan membandingkan model ini dan beberapa model lain, menunjukkan bahwa model NARX memberikan akurasi tertinggi dalam 6 kasus pada percobaan pertama dan mampu meningkatkan akurasi model. (Du, et al. 2018) menggunakan model ini pada penelitiannya dan menghasilkan kinerja peramalan jaringan yang baik, menunjukkan bahwa model NARX yang digunakan memiliki kemampuan generalisasi yang baik. (Liang, et al. 2021) dalam penelitiannya menunjukkan bahwa, untuk prakiraan 6 jam dan 24 jam, akurasi dan stabilitas model NARX secara signifikan lebih baik daripada tiga model yang umum digunakan, terutama dalam prediksi jam atom sesium satelit.

Model *Vector Autoregression* (VAR) adalah metode statistik di mana setiap variabel diperlakukan secara simetris dan nilai pada waktu ke-*t* dapat dipengaruhi oleh nilai saat ini dan nilai pada masa lalu (Lütkepohl 2005). *Vector Autoregressive* (VAR) dapat menggambarkan perilaku dinamis dari

data menggunakan hubungan antara beberapa variabel. (Liu, Roberts and Sioshansi 2018) menggunakan model ini dalam memprediksi cuaca untuk permintaan maupun penawaran (*supply and demand*) sistem kelistrikan, menunjukkan bahwa VAR memberikan nilai skor keterampilan lebih dari dua kali lipat dari skor keterampilan model prediksi lain dalam literatur, model juga menyediakan prediksi yang mengungguli prakiraan persistensi antara 6% dan 80% dalam hal kesalahan absolut rata-rata (MAE). (Ankamah, Nokoe and Iddrisu 2018) melakukan pemodelan tren visibilitas iklim dan malaria di Ghana, menunjukkan bahwa malaria dipengaruhi oleh tiga variabel iklim dengan MAPE sebesar 12,96% dengan nilai $p = 14$ yang merupakan prakiraan yang berpotensi sangat baik.

Berdasarkan penelitian yang melibatkan model *Recurrent NARX Neural Networks* (Siegelmann, Horne and Giles 1997) untuk membuktikan kapabilitas kemampuan model ini jika dibandingkan dengan model *fully-connected* milik *Recurrent Neural Network* (RNN), serta kemampuan VAR dalam memahami dan menggunakan hubungan antara beberapa variabel untuk menggambarkan perilaku dinamis dari data, maka penulis tertarik untuk melakukan penelitian dengan judul “Perbandingan *Recurrent NARX Neural Networks* dan *Vector Autoregressive* (VAR) dalam Memprediksi Kualitas Udara DKI Jakarta” untuk mengimplementasikan model serta mengetahui nilai presisi dan *recall* yang dihasilkan apabila digunakan model *Recurrent NARX Neural Networks* maupun *Vector Autoregressive* (VAR).

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka penulis membuat rumusan masalah sebagai berikut:

1. Bagaimana langkah untuk mengimplementasikan model *Recurrent NARX Neural Networks* dan *Vector Autoregressive* (VAR) dalam memprediksikan kualitas udara DKI Jakarta?
2. Bagaimana nilai presisi dan *recall* yang dihasilkan dari pengimplementasian model *Recurrent NARX Neural Networks* dan

Vector Autoregressive (VAR) dalam memprediksikan kualitas udara DKI Jakarta?

C. Batasan Masalah

Terdapat beberapa batasan masalah dalam pelaksanaan penelitian ini, di antaranya:

1. Data yang digunakan dalam penelitian ini adalah data Indeks Standar Pencemar Udara (ISPU) DKI Jakarta pada rentang Januari 2018 hingga Desember 2021, yaitu sebanyak 1.461 baris.
2. Sebagai perhitungan, akan digunakan beberapa parameter Indeks Standar Pencemar Udara (ISPU), yaitu partikular (PM10), karbon monoksida (CO), sulfur dioksida (SO2), nitrogen dioksida (NO2), dan ozon (O3).
3. Jenis arsitektur NARX yang akan digunakan adalah model seri-paralel (NARX-S) (Narendra and Parthasarathy 1990) dengan satu lapisan tersembunyi.
4. Model akan dibuat berdasarkan lima lokasi Stasiun Pemantau Kualitas Udara (SPKU) yang ada di Provinsi DKI Jakarta, yaitu DKI1 yang berlokasi di Bundaran HI, DKI2 yang berlokasi di Kelapa Gading, DKI3 yang berlokasi di Jagakarsa, DKI4 yang berlokasi di Lubang Buaya, dan DKI5 yang berlokasi di Kebon Jeruk.
5. Penelitian ini akan menggunakan bahasa pemrograman *python* dengan menggunakan beberapa *library* dasar yang akan dibutuhkan, seperti *pandas* dan *statsmodels* untuk pengolahan data, *numpy* untuk membantu pembuatan model, serta *sklearn.metrics* untuk mengevaluasi model.
6. Untuk memaksimalkan visualisasi dalam proses pengolahan data akan digunakan *Google Colab* yang merupakan lingkungan (*environment*) *Jupyter* yang berjalan di *cloud* sehingga dapat dijalankan cukup melalui *browser* saja. Sedangkan dalam proses implementasi model akan digunakan *Visual Studio Code* yang merupakan lingkungan

pengembangan terintegrasi (IDE) yang mendukung banyak fitur dalam pengembangan perangkat lunak.

7. Untuk menilai kemampuan model dalam memprediksikan hasil dari setiap data, dalam penelitian ini akan digunakan presisi dan *recall*.

D. Tujuan Penelitian

Tujuan yang ingin dicapai oleh penulis dalam pelaksanaan penelitian ini adalah untuk mengimplementasikan model *Recurrent NARX Neural Networks* dan *Vector Autoregressive (VAR)* dalam memprediksikan kualitas udara DKI Jakarta serta mengetahui nilai presisi dan *recall* yang dihasilkan.

E. Manfaat Penelitian

Manfaat yang diharapkan oleh penulis dalam pelaksanaan penelitian ini adalah:

1. Bagi Penulis

Penulis dapat menambah wawasan mengenai pengimplementasian model *Recurrent NARX Neural Networks* dan *Vector Autoregressive (VAR)* untuk membantu menyelesaikan permasalahan yang sering terjadi di masyarakat.

2. Bagi Program Studi Ilmu Komputer Universitas Negeri Jakarta

Penelitian ini dapat memberikan gambaran bagi seluruh mahasiswa Program Studi Ilmu Komputer Universitas Negeri Jakarta tentang proses memprediksikan kualitas udara DKI Jakarta menggunakan model *Recurrent NARX Neural Networks* dan *Vector Autoregressive (VAR)*.