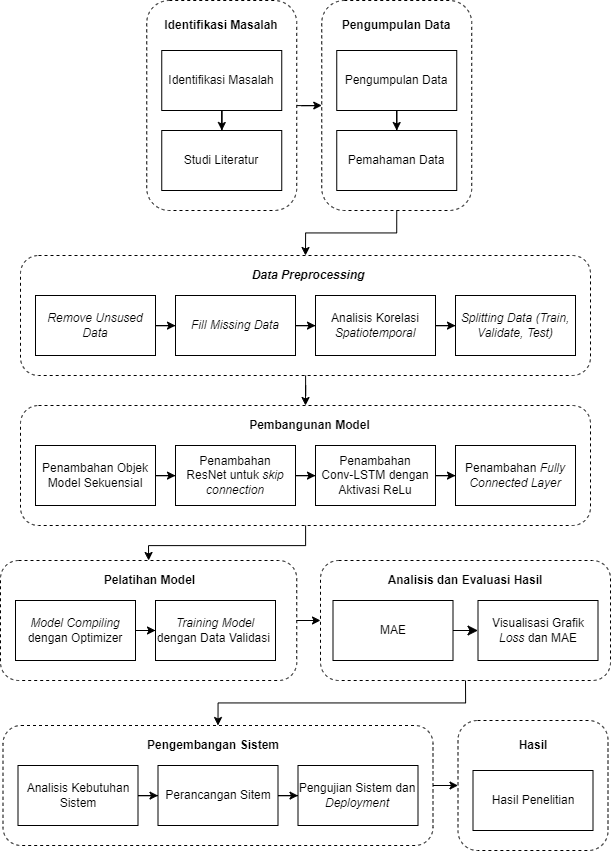
# **BAB III**

# **METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM**

# **Metodologi Penelitian**

Pada bagian ini akan dibahas mengenai metodologi yang akan dilakukan dalam penelitian ini, yaitu mengatasi *vanishing gradient* menggunakan ResNet Conv-LSTM pada prediksi konsentrasi polutan dalam kualitas udara. Penelitian ini menerapkan metode kuantitatif menggunakan data numerik dan statistik untuk menafsirkan informasi yang digunakan dalam pengujian hipotesis dengan teori-teori yang sudah ada. Adapun data yang digunakan dalam penelitian ini adalah data historis zat polutan dan meteorologi Provinsi DKI Jakarta. Tahapan penelitian ditunjukkan pada Gambar x



Tahapan penelitian di mulai dari identifikasi masalah, pengumpulan data, *data preprocessing*, pembangunan model, pelatihan model, analisis dan evaluasi hasil, pengembangan sistem, dan diakhir dengan hasil penelitian.

# **Identifikasi Masalah**

1. Identifikasi Masalah

Penelitian ini diawali dengan mengidentifikasi masalah yang akan diangkat dan diselesaikan dalam penelitian. Masalah yang diangkat adalah masalah *vanishing gradient* akibat meningkatnya kompleksitas jaringan pada prediksi data *spatiotemporal* yang menyebabkan akurasi menjadi tidak optimal.

1. Studi Literatur

Studi literatur dilakukan untuk memperoleh informasi yang relevan yang akan digunakan sebagai dasar atau acuan pada penelitian ini. Tahapan ini dilakukan dengan menelusuri beberapa sumber seperti buku, jurnal, dan penelitian terdahulu mengenai *vanishing gradient* dan prediksi konsentrasi polutan dalam kualitas udara. Informasi yang didapatkan digunakan sebagai dasar untuk menyelesaikan masalah dan mencapai tujuan penelitian. Informasi dari penelitian-penelitian terdahulu dapat dilihat dalam *State of the Art* pada **Tabel 1**.

# **Pengumpulan Data**

1. Pengumpulan Data

Penelitian ini menggunakan kombinasi data primer-sekunder, yaitu data historis konsentrasi polutan dari lima Stasiun Pemantauan Kualitas Udara di DKI Jakarta dan data meteorologi dari Stasiun Meteorologi di Jakarta Pusat, DKI Jakarta. Masing-masing data diperoleh dari *website* Satu Data Jakarta untuk data konsentrasi polutan dan Data Online Pusat Database BMKG untuk data meteorologi dalam bentuk *excel*. Data ini dikumpulkan setiap 1 hari dari 1 Januari 2019 – 31 Desember 2021 sebanyak 43.480 data.

Pada penelitian ini, dipilih lima konsentrasi polutan, yaitu PM10, SO2, CO, O3, dan NO2 yang diambil dari Jakarta Pusat, Jakarta Barat, Jakarta Utara, Jakarta Selatan, dan Jakarta Timur serta lima faktor meteorologi, yaitu suhu rata-rata, kelembaban rata-rata, kecepatan angin rata-rata, curah hujan, dan arah angin saat kecepatan maksimum.

1. Pemahaman Data

Data konsentrasi polutan dan meteorologi yang diperoleh memiliki beberapa parameter yang digunakan untuk melakukan prediksi menggunakan metode ResNet Conv-LSTM. Detail dari setiap parameter dapat dilihat pada **Tabel**.

| **No** | **Variabel** | **Deskripsi** | **Keterangan** |
| --- | --- | --- | --- |
| **I** | **Data Konsentrasi Polutan** | | |
| 1 | tanggal | Tanggal pengukuran konsentrasi polutan | Format penanggalan mm/dd/yyyy |
| 2 | stasiun | Nomor/nama Stasiun Pemantauan Kualitas Udara (SPKU) di DKI Jakarta | Sebaran SPKU:   1. DKI 1 – Bundaran HI 2. DKI 2 – Kelapa Gading 3. DKI 3 – Jagakarsa 4. DKI 4 – Lubang Buaya 5. DKI 5 – Kebon Jeruk |
| 3 | pm10 | Rata-rata konsentrasi zat PM10 dalam periode waktu 1 hari pengukuran | Data yang tidak ada dikosongkan atau diberi tanda --- |
| 4 | so2 | Rata-rata konsentrasi zat SO2 dalam periode waktu 1 hari pengukuran | Data yang tidak ada dikosongkan atau diberi tanda --- |
| 5 | co | Rata-rata konsentrasi zat CO dalam periode waktu 1 hari pengukuran | Data yang tidak ada dikosongkan atau diberi tanda --- |
| 6 | no2 | Rata-rata konsentrasi zat NO2 dalam periode waktu 1 hari pengukuran | Data yang tidak ada dikosongkan atau diberi tanda --- |
| 7 | max | Nilai konsentrasi polutan maksimum dalam 1 hari | Data yang tidak ada diberi nilai 0 |
| 8 | critical | Zat polutan dengan nilai konsentrasi maksimum dalam 1 hari | Data yang tidak ada dikosongkan |
| 9 | categori | Kategori Indeks Standar Pencemar Udara (ISPU) dalam 1 hari | Kategori ISPU:   1. BAIK 2. SEDANG 3. TIDAK SEHAT 4. SANGAT TIDAK SEHAT 5. TIDAK ADA DATA |
| **II** | **Data Meteorologi** | | |
| 1 | Tanggal | Tanggal pengukuran faktor meterologi | Format penanggalan dd-mm-yyy |
| 2 | Tavg | Nilai suhu rata-rata dalam waktu 1 hari pengukuran | Data yang tidak terukur diberi nilai 8888, sedangkan data yang tidak ada diberi nilai 9999 atau dikosongkan. |
| 3 | RH\_avg | Nilai kelembaban rata-rata dalam waktu 1 hari pengukuran | Data yang tidak terukur diberi nilai 8888, sedangkan data yang tidak ada diberi nilai 9999 atau dikosongkan. |
| 4 | RR | Nilai curah hujan dalam waktu 1 hari pengukuran | Data yang tidak terukur diberi nilai 8888, sedangkan data yang tidak ada diberi nilai 9999 atau dikosongkan. |
| 5 | ddd\_x | Nilai arah angin dalam kecepatan maksimum dalam waktu 1 hari pengukuran | Data yang tidak terukur diberi nilai 8888, sedangkan data yang tidak ada diberi nilai 9999 atau dikosongkan. |
| 6 | ff\_avg | Nilai kecepatan angin rata-rata dalam waktu 1 hari pengukuran | Data yang tidak terukur diberi nilai 8888, sedangkan data yang tidak ada diberi nilai 9999 atau dikosongkan. |

# ***Data Preprocessing***

Data yang telah didapatkan kemudian diproses agar menjadi lebih terstruktur sehingga dapat digunakan dalam pembuatan dan pelatihan model. Beberapa proses yang dilakukan pada tahapan ini, yaitu menghapus data-data yang tidak diperlukan, mengisi nilai yang hilang, menganalisis analisis korelasi *spatiotemporal*, dan membagi data.

1. *Remove Unused Data*

Penghapusan data dilakukan karena tidak semua variabel pada *dataset* akan digunakan dalam pembuatan model. Variabel yang dihapus meliputi max, critical, dan categori pada *dataset* konsentrasi polutan.

1. *Fill Missing Value*

Setelah variabel yang tidak digunakan telah dihapus, selanjutnya adalah mengisi nilai yang hilang pada *dataset* konsentrasi polutan dan meteorologi. Pengisian nilai yang hilang ini menggunakan metode interpolasi linear. Misalnya, pada *dataset* terdapat nilai yang hilang pada tanggal 27 Januari 2019 seperti pada **Tabel**.

|  |  |  |
| --- | --- | --- |
| **Tanggal** | **CO** | **O3** |
| 1/26/2019 | 33 | 118 |
| 1/27/2019 | 14 |  |
| 1/28/2019 | 22 | 64 |

Di, mana dan , maka penerapan metode interpolasi linear menggunakan rumus (2.17) adalah sebagai berikut:

(2.17)

(3.1)

(3.2)

(3.3)

(3.4)

Sehingga, diperoleh hasil yang merupakan nilai O3 pada tanggal 27 Januari 2019 sebesar 137.

1. Analisis Korelasi *Spatiotemporal*
2. *Data Splitting*

Tahapan ini diakhiri dengan melakukan *data splitting* untuk membagi data menjadi tiga bagian, yaitu *data train, validate,* dan *test*. Pembagian data yaitu 70% untuk *data* *train*, 15% untuk *data* *validate*, dan 15% untuk *data test*.

# **Pembangunan Model**

# **Pelatihan Model**

# **Analisis dan Evaluasi Hasil**

# **Pengembangan Sistem**

# **Hasil**

# **Pengembangan Sistem**