

BAB 3

METODE PENELITIAN

Metode penelitian ini menggunakan pendekatan *Machine Learning* yang bertujuan untuk mengembangkan model prediksi cuaca dengan menggunakan data historis dari BMKG Yogyakarta. Penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan data, preprocessing data, forecasting, serta evaluasi dan pengujian hasil prediksi. Setiap tahapan penelitian dirancang agar model yang dibangun memiliki akurasi tinggi dalam memprediksi kondisi cuaca, khususnya dalam mengenali pola cuaca ekstrem dan tren jangka panjang.

3.1 BAHAN DAN ALAT PENELITIAN

Penelitian ini menggunakan data historis cuaca Yogyakarta periode 2020-2024 dari Stasiun Klimatologi D.I Yogyakarta (BMKG), mencakup variabel suhu, curah hujan, kelembaban, dan kecepatan angin dalam format CSV/Excel yang sesuai untuk analisis berbasis *Machine Learning (ML)*.

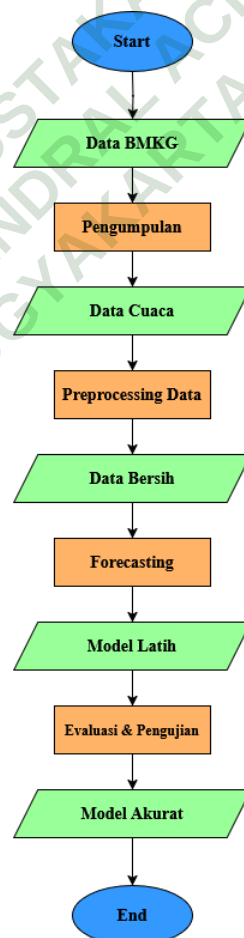
Untuk mendukung pengolahan data dan pelatihan model prediksi cuaca, diperlukan sejumlah alat yang mencakup sistem operasi, perangkat keras, dan perangkat lunak sebagai berikut:

1. Sistem Operasi Minimum:
 - Windows 10/11.
 - Linux (Ubuntu/Debian).
2. Perangkat Keras:
 - a. Komputer/Laptop dengan spesifikasi minimal:
 - Prosesor: Intel Core i3 atau lebih tinggi.
 - RAM: 8GB.
 - Penyimpanan: 256GB SSD.
3. Perangkat Lunak:
 - a. Python dengan library:
 - Pandas (1.5.3).

- NumPy (1.23.5).
 - Scikit-Learn (1.1.3).
 - TensorFlow (2.10.0).
 - Statsmodels (0.13.2).
 - Streamlit (1.25.0).
- b. *Google Colaboratory* atau *Jupyter Notebook* untuk pelatihan model.
- c. Aplikasi pengolah kata dan *spreadsheet* sebagai pendukung dokumentasi.

3.2 JALAN PENELITIAN

Pada Gambar 3.1 menunjukkan alur dari penelitian yang mencakup tahapan dari pengumpulan data hingga evaluasi model prediksi cuaca.



Gambar 3.1 *Flowchart* Jalan Penelitian

Pada tahapan penelitian yang ditunjukkan pada Gambar 3.1 menjelaskan proses dalam membangun dan mengevaluasi model prediksi cuaca berbasis *Machine Learning (ML)*. Berikut adalah langkah-langkahnya:

1. Pengumpulan Data

- a. Proses: Mengumpulkan data historis cuaca dari BMKG melalui Stasiun Klimatologi D.I Yogyakarta untuk periode 2020-2024.
- b. Data yang dikumpulkan:
 - Suhu.
 - Curah hujan.
 - Kelembaban.
 - Kecepatan angin.
- c. Hasil: Dataset dalam format CSV/Excel yang siap untuk diproses lebih lanjut.

2. Preprocessing Data

- a. Proses:
 - 1) Membersihkan data dari duplikasi atau inkonsistensi.
 - 2) Mendeteksi dan menangani *outlier* menggunakan metode *Interquartile Range (IQR)*.
 - *Outlier* dideteksi dengan menghitung batas bawah ($Q1 - 1.5 \times IQR$) dan batas atas ($Q3 + 1.5 \times IQR$).
 - Nilai yang berada di luar rentang dianggap sebagai *outlier*.
 - Menggunakan visualisasi boxplot untuk memperjelas deteksi pada setiap variabel cuaca seperti suhu, curah hujan, kelembaban, dan kecepatan angin. Boxplot digunakan karena dapat membantu untuk melihat sebaran data dan posisi pencilan secara visual.
 - 3) Menghapus data yang hilang atau tidak valid.
 - 4) Melakukan normalisasi data agar memiliki skala yang seragam.
 - 5) Transformasi data ke dalam format *time series* agar sesuai model *Machine Learning (ML)*.

b. Hasil: Dataset yang bersih dan siap digunakan untuk pemodelan.

3. Forecasting

a. Proses:

1) Pemilihan Model: Menentukan algoritma yang sesuai untuk prediksi cuaca. Dengan opsi model sebagai berikut:

- *Long Short-Term Memory (LSTM)*: Cocok untuk memodelkan data deret waktu karena mampu mengingat pola jangka panjang.
- *Random Forest (RF)*: Algoritma berbasis *ensemble* yang dapat mengenali pola variabel lingkungan untuk estimasi curah hujan.
- *XGBoost*: Pengembangan dari *gradient boosting* yang efektif untuk meningkatkan akurasi pada data numerik kompleks.
- *AutoRegressive Integrated Moving Average (ARIMA)*: Model statistik klasik yang digunakan untuk meramalkan tren jangka panjang dalam data curah hujan.

Hasil: Model terbaik dipilih berdasarkan evaluasi performa menggunakan metrik RMSE, MAPE, dan R^2 Score.

2) Training Model: Model yang terpilih akan dilatih dengan tahapan sebagai berikut:

- Pembagian Data: Data historis dari tahun 2020 hingga 2024 dibagi menjadi dua bagian, yaitu 80% untuk pelatihan (training) dan 20% untuk pengujian (testing). Namun dalam pendekatan utama, digunakan strategi pembagian berdasarkan waktu, yaitu data tahun 2020–2023 digunakan untuk pelatihan, dan data tahun 2024 untuk pengujian, agar sesuai dengan skenario prediksi jangka panjang.
- Proses Pelatihan: Model dilatih menggunakan data historis curah hujan (RR), suhu rata-rata, kelembapan

rata-rata, dan kecepatan angin rata-rata. Tujuan pelatihan adalah agar model mampu mengenali pola cuaca dari masa lalu dan menghasilkan prediksi curah hujan harian secara akurat.

- Penyimpanan Model: Model yang telah selesai dilatih disimpan (persistensi model) untuk proses evaluasi, prediksi ulang, dan penerapan ke dalam dashboard interaktif.

b. Hasil: Model *Machine Learning (ML)* yang telah terlatih dan siap digunakan untuk *forecasting* cuaca.

4. Evaluasi dan Pengujian Model

a. Proses:

- 1) Model diuji menggunakan data pengujian (*testing set*) yang tidak digunakan selama proses pelatihan (*training*).
- 2) Dilakukan optimasi model, apabila diperlukan, menggunakan teknik seperti *Grid Search* untuk mencari kombinasi parameter terbaik.
- 3) Evaluasi performa model dilakukan dengan menggunakan beberapa metrik berikut:

- *Root Mean Squared Error (RMSE)*: Mengukur rata-rata selisih kuadrat antara nilai prediksi dan nilai aktual.
- *Mean Absolute Percentage Error (MAPE)*: Mengukur rata-rata persentase kesalahan dari prediksi terhadap data aktual.
- *R² Score (Koefisien Determinasi)*: Menunjukkan seberapa besar variasi data yang dapat dijelaskan oleh model.

4) Hasil prediksi model dibandingkan dengan:

- Data aktual dari BMKG, sebagai dasar evaluasi performa model.
- Prediksi cuaca dari BMKG, untuk membandingkan tingkat akurasi antara hasil model penelitian dan metode

prediksi resmi yang digunakan oleh BMKG, seperti Numerical Weather Prediction (NWP).

- b. Hasil: Model terbaik dipilih berdasarkan evaluasi terhadap data aktual dan, jika memungkinkan, dibandingkan dengan data aktual dari BMKG.

3.3 OBSERVASI DAN WAWANCARA PENDUKUNG

Dalam rangka menunjang pengumpulan data dan pemahaman terhadap mekanisme pengamatan serta pengolahan data cuaca oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), penulis melakukan observasi lapangan dan wawancara langsung dengan petugas BMKG. Kegiatan ini dilakukan di dua unit, yaitu Stasiun Klimatologi dan Stasiun Meteorologi BMKG Yogyakarta, yang masing-masing memiliki peran berbeda dalam proses pengamatan dan prediksi cuaca. Observasi dan wawancara dilakukan untuk memperoleh informasi terkait proses pengamatan, pengumpulan, pengolahan data cuaca, serta perbedaan peran antar unit dalam hal prediksi cuaca jangka pendek dan analisis iklim jangka panjang.

Tabel 3.1 Ringkasan Wawancara 1 – Stasiun Klimatologi Yogyakarta

ASPEK	ISI/PENJELASAN
Topik	Proses pengamatan dan pencatatan data cuaca harian
Waktu dan Tempat	Hari ke-1, Stasiun Klimatologi BMKG
Nama Narasumber	Pak Subandi
Peran Stasiun	Fokus pada pengamatan dan pencatatan data harian, bukan prediksi.
Data Diamati	Suhu, kelembaban, curah hujan, tekanan udara, angin, penguapan, awan, visibilitas
Sistem Pengamatan	Terdapat alat manual dan otomatis, data dicatat setiap jam (khususnya jam 07.00).

Tujuan Data	Dikirim ke pusat untuk keperluan prakiraan dan dokumentasi.
Keterangan Tambahan	Data harian tidak langsung digunakan untuk prediksi di unit ini. Prediksi jangka pendek dilakukan oleh bagian Meteorologi

Tabel 3.2 Ringkasan Wawancara 2 – Stasiun Klimatologi Yogyakarta

ASPEK	ISI/PENJELASAN
Topik	Proses pengolahan dan pelaporan data cuaca
Waktu dan Tempat	Hari ke-2, Bagian Pengolahan Data BMKG
Nama Narasumber	Bu Gesti
Peran Unit	Fokus pada pengolahan data pengamatan menjadi laporan iklim.
Format Data	Data diolah berdasarkan juknis F-Klim 71 menjadi laporan bulanan dan dasarian.
Komponen Pengolahan	Suhu rata-rata, suhu maksimum dan minimum, curah hujan, penyinaran matahari, serta cuaca signifikan.
Metode Perhitungan	Rata-rata suhu: $(2 \times \text{Jam 7} + \text{Jam 13} + \text{Jam 18})/4$, curah hujan harian dari pukul 07.00–07.00 keesokan hari.
Output Pengolahan	Dikirim ke unit pusat atau stakeholder, digunakan untuk perbandingan atau analisis iklim.
Keterangan Tambahan	Unit ini tidak menghasilkan grafik atau visualisasi prediksi; fokus hanya pada konsolidasi data dan laporan rutin.

Tabel 3.3 Ringkasan Wawancara 3 – Stasiun Meteorologi Yogyakarta

ASPEK	ISI/PENJELASAN
Topik	Sistem prediksi cuaca jangka pendek dan kendala pengambilan data
Waktu dan Tempat	Hari ke-3, Stasiun Meteorologi BMKG
Nama Narasumber	Kepala Stasiun Meteorologi Yogyakarta
Peran Stasiun	Fokus pada prediksi cuaca jangka pendek (1–7 hari ke depan).
Sistem Prediksi	Menggunakan komponen kompleks: tekanan udara, suhu, kelembaban, angin, konvergensi/divergensi, dan lifting index.
Data Prediksi	Dikumpulkan dan dianalisis per jam, diproses menggunakan sistem internal BMKG.
Kendala ML	Data prediksi cuaca pendek sangat kompleks, sulit diakses publik. Tidak disarankan menggunakan data terbatas untuk prediksi harian berbasis ML.
Saran Narasumber	Gunakan data klimatologi dari BMKG untuk prediksi jangka panjang dan model ML yang sederhana seperti XGBoost atau Random Forest.

Dari hasil observasi dan wawancara di kedua bagian BMKG tersebut, diperoleh pemahaman bahwa terdapat pemisahan tugas dan spesialisasi antara unit pengamatan, pengolahan, dan prediksi. Data dari Stasiun Klimatologi digunakan untuk mencatat kondisi cuaca harian dan menjadi dasar dalam laporan iklim. Unit pengolahan data kemudian menyusun laporan secara periodik berdasarkan petunjuk teknis (juknis) tertentu, sementara prediksi jangka pendek dilakukan oleh Stasiun Meteorologi menggunakan metode yang kompleks dan sistem tertutup.

Berdasarkan wawancara ini, pendekatan yang paling memungkinkan untuk penelitian prediksi curah hujan adalah menggunakan data klimatologi harian sebagai dasar pelatihan model *Machine Learning* jangka panjang. Data tersebut memiliki struktur yang cukup lengkap (TN, TX, TAVG, RH_AVG, RR, FF_X, FF_AVG) dan tersedia secara publik, sehingga dapat digunakan secara objektif tanpa melibatkan sistem prediksi internal BMKG.

PERPUSTAKAAN
JENDRAL ACHMAD YANI
YOGYAKARTA