

BAB III METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian ini bersifat *kuantitatif* atau memiliki metode penelitian *kuantitatif*. Karena pada jenis penelitian ini menggunakan angka-angka atau mengumpulkan data yang mendalam dan lengkap untuk menggambarkan fenomena atau fakta yang diteliti. Dalam penelitian ini, kelengkapan dan kedalaman data penelitian sangat penting. Studi ini mengumpulkan data dengan mengambil data dari platform Kaggle yang diperoleh dalam format .xlsx atau .csv.

3.2 Tahap Penelitian

Tahapan Penelitian yang akan dilakukan seperti berikut:



Gambar 3.2.1 Tahapan Penelitian

Penjelasan terkait tahapan dalam riset berikut bisa dicermati dalam Gambar 3.2.1, serta penjelasan dari gambar tersebut dapat dijelaskan sebagai berikut:

3.3 Identifikasi Masalah

Identifikasi permasalahan pada riset berikut yakni memprediksi suhu cuaca di Nowergia menggunakan metode Algoritma *Long Short-Term Memory*.

3.4 Pengumpulan Data

Metode pengumpulan data untuk riset berikut ialah merupakan data real

yang didapatkan dari platform Kaggle, bagian Data yang di kumpulkan berupa data temperatur cuaca di Nowwergia, yang menyajikan data temperatur cuaca di Nowwergia pada tahun 2015 sampai dengan 2019, dengan data yang di peroleh sebanyak 1.826 data.

3.5 Pengolahan Data (Preprocessing)

Preprocessing adalah langkah membersihkan data mentah yang berantakan dan mengubahnya menjadi data yang terstruktur dengan baik [15]. Menganalisis data yang tidak difilter dapat menyebabkan masalah dengan hasil akhir, karena tidak sesuai dengan ekspektasi pemrosesan data. Oleh karena itu, penyajian dan kualitas data sangat penting sebelum melakukan analisis apa pun. Data preprosesing antara lain : [25]

1. Data Reduction

Pada tahap ini data yang terkumpul kemudian dirangkum dan diseleksi agar dapat dilakukan penentuan data, pemusatan perhatian dalam penyederhanaan data, abstraksi data agar dapat sesuai dengan data yang dibutuhkan agar dapat diproses sesuai dengan metode yang digunakan, dalam proses ini akan menghilangkan data atribut yang kurang penting, dimaksudkan untuk meminimalkan datanya dan tetap menjaga ketepatannya [16].

2. Data Cleaning

Ada beragam bagian yang tidak berkaitan serta hilang dalam data riset berikut. Guna menutupi bagian ini, dijalankan cleaning data, cleaning data perihal data yang hilang, data yang berisik, dan transformasi data. Bila terjadi kehilangan data, cleaning dijalankan guna meminimalisir ketidaksesuaian data. Dalam riset berikut langkah selanjutnya ialah mencari nilai data yang hilang guna meminimalisir nilai data yang hilang [17].

3. Data Transformation

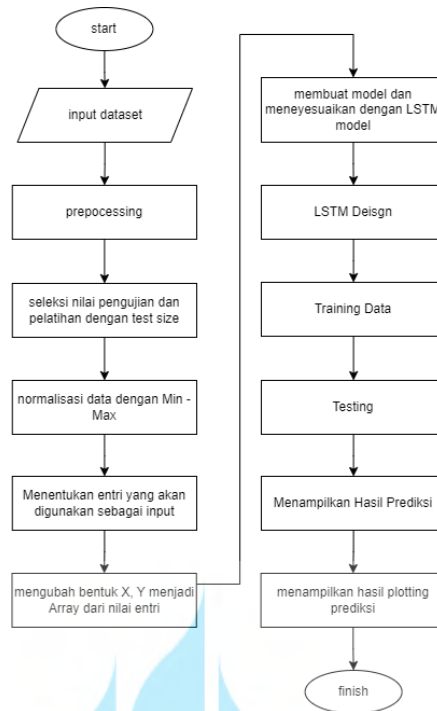
Dengan konversi data, data diubah dalam format yang selaras. Normalisasi ialah tahapan penskalaan nilai data pada rentang yang ditentukan guna menentukan tidak ddiapati data berlebih [26]. pada tahap *transform* data melakukan tahap pendekatan *transformation* dengan menggunakan 2 tahap yaitu *Attribute Construction* dan juga *agregation*, dan penjelasan 2 tahap tersebut antara lain :

- a. *Agregation* pada tahap pendekatan ini melakukan pengelompokan data dengan merubah bentuk data tanpa makna dari data yang di gunakan, yaitu melakukan pengelompokan data pada bulan, tanggal, dan tahun agar dapat mempermudah dalam memproses dan tetap dapat di proses.
- b. *Attribute Construction* pada tahap pendekatan ini melakukan pembentukan *attribute & variable* baru yang sudah ada untuk membantu meningkat kan akurasi dan pemahaman yaitu menambahkan atribut tahun pada *attribute* tanggal pada data penelitian ini agar dapat diporses.

3.6 Implementasi Algoritma Long Short Term Memory

Dari tiap proses pada gambar 3.6.1 dapat dijelaskan sebagai berikut yaitu setelah inputan dataset sampai dengan mengubah bentuk X dan Y menjadi array maka akan di buat model mensklalakan ulang nilai *train* dan *test*, selanjutnya membentuk *LSTM design*, dan melakukan *training* dan *testing* dataset, selanjutnya akan menghitung jumlah *Root Mean Squared Error* (RMSE), selanjutnya akan membalikan prediksi ke proses pembuatan dan penyesuaian model LSTM untuk memastikan apa yang di prediksi sama dengan data asli, selanjutnya melakukan *ploting* hasil prediksi, dan setelah itu akan dihasilkan out put berupa *ploting* prediksi.

UNIVERSITAS
MERCU BUANA



Gambar 3.6.1 Flowchart Algoritma Long Short Term Memory

Penjelasan dari gambar tersebut dapat dijelaskan sebagai berikut:

1. Input dataset ialah kumpulan data yang digunakan sebagai dasar untuk analisis dan penelitian dalam skripsi ini. Dataset ini berisi informasi tentang fluktuasi temperatur cuaca di Norwegia dari tahun 2015 hingga 2019, dengan fokus pada kolom "tmin" (temperatur minimum). Data tersebut mencakup temperatur minimum harian di berbagai lokasi di Norwegia selama periode tersebut. Dataset ini akan digunakan untuk analisis dan visualisasi fluktuasi temperatur cuaca guna mendapatkan wawasan tentang perubahan iklim di Norwegia.
2. Preprocessing atau pra-pemrosesan data adalah serangkaian langkah yang dilakukan untuk mempersiapkan data mentah sebelum diproses dalam analisis atau pemodelan. Tujuan dari preprocessing adalah untuk meningkatkan kualitas data, menghilangkan kecacatan atau noise, dan mengubah data menjadi bentuk yang lebih sesuai untuk analisis.
3. Normalisasi dengan metode Min-Max ialah satu diantara teknik pra-pemrosesan data yang digunakan untuk mengubah rentang nilai data menjadi rentang yang spesifik, sering kali antara 0 dan 1[19]. Tujuannya adalah untuk membawa semua nilai dalam dataset ke skala yang seragam

dan mempertahankan proporsi relatif antara nilai-nilai tersebut.

4. Menentukan entri yang akan digunakan sebagai input dalam analisis atau pemodelan data melibatkan pemilihan variabel atau fitur yang dianggap memiliki hubungan atau pengaruh terhadap variabel target atau output yang ingin diprediksi.
5. Untuk mengubah bentuk X dan Y menjadi array dari nilai entri, Anda dapat menggunakan pustaka Python seperti NumPy atau Pandas. Dengan NumPy, Anda dapat mengonversi nilai entri X dan Y menjadi array NumPy menggunakan fungsi **np.array()**. Sedangkan dengan Pandas, Anda dapat mengubah nilai entri X menjadi DataFrame Pandas dan Y menjadi Series Pandas. Mengubah nilai entri menjadi array memungkinkan Anda untuk memanipulasi dan menganalisis data dengan lebih mudah menggunakan fitur-fitur yang disediakan oleh pustaka-pustaka tersebut.
6. Untuk membuat dan menyesuaikan model LSTM, langkah-langkahnya adalah sebagai berikut. Pertama, impor pustaka yang dibutuhkan seperti TensorFlow atau Keras. Selanjutnya, persiapkan data dalam format sekuensial yang sesuai, dengan membagi data menjadi fitur (X) dan target (Y). Kemudian, bangun arsitektur model LSTM dengan menambahkan lapisan LSTM, Dense, dan lainnya sesuai kebutuhan. Setelah itu, kompilasi model dengan menentukan fungsi loss, optimizer, dan metrik evaluasi. Lakukan pelatihan model dengan menggunakan data pelatihan, dengan menentukan jumlah epoch dan ukuran batch yang sesuai. Setelah melatih model, evaluasi kinerjanya menggunakan data pengujian. Jika perlu, lakukan penyesuaian parameter dan ulangi proses pelatihan. Setelah model dilatih dan dievaluasi, Anda dapat menggunakannya untuk memprediksi nilai target baru atau menerapkan fungsi yang diinginkan.
7. Desain LSTM ialah suatu arsitektur yang dipakai untuk memodelkan dan menganalisis data sekuensial dengan kemampuan mengingat informasi jangka panjang dan jangka pendek. LSTM dirancang dengan menggunakan lapisan memori khusus yang memungkinkan pengolahan data sekuensial dengan konteks jangka panjang yang lebih baik. Arsitektur

LSTM terdiri dari lapisan LSTM, yang memiliki unit memori khusus yang disebut "cell state". Cell state ini memungkinkan model untuk mengingat informasi penting dari masa lalu dan mengabaikan informasi yang tidak relevan. Selain itu, arsitektur LSTM juga menggunakan gate, seperti gate input, gate forget, dan gate output, yang mengatur aliran informasi dalam dan keluar dari unit LSTM. Desain LSTM ini memungkinkan model untuk mengatasi masalah ketergantungan jarak jauh dan mempelajari pola dan relasi yang kompleks dalam data sekuensial.

8. Training data merujuk pada data yang digunakan untuk melatih atau mengajari model atau algoritma dalam konteks pembelajaran mesin. Data pelatihan merupakan dataset yang berisi pasangan input dan output yang sudah diketahui, yang digunakan untuk mengenalkan model pada pola dan hubungan dalam data.
9. Testing (pengujian) merujuk pada proses menguji kinerja model atau algoritma pada data yang tidak digunakan selama proses pelatihan. Tujuan dari pengujian adalah untuk mengevaluasi sejauh mana model yang dilatih dapat menghasilkan prediksi yang akurat dan umum pada data baru yang belum pernah dilihat sebelumnya.
10. Menampilkan hasil prediksi adalah proses untuk memperlihatkan nilai-nilai yang diprediksi oleh model pada data yang belum pernah dilihat sebelumnya. Hal ini dilakukan untuk mengevaluasi performa model dan memahami sejauh mana model dapat membuat prediksi yang akurat dan berguna. Hasil prediksi dapat ditampilkan dalam berbagai cara, seperti tabel atau daftar yang memuat nilai prediksi bersama dengan nilai asli, grafik atau plot untuk memvisualisasikan perbedaan antara nilai prediksi dan nilai asli, serta metrik evaluasi yang memberikan angka atau grafik yang menggambarkan kinerja model secara numerik. Menampilkan hasil prediksi dengan cara yang jelas dan informatif membantu pemahaman tentang kualitas prediksi model dan memungkinkan pengambilan keputusan yang tepat berdasarkan prediksi tersebut.
11. Menampilkan hasil plotting prediksi melibatkan visualisasi grafik atau plot yang menggambarkan perbandingan antara nilai prediksi yang dihasilkan

oleh model dengan nilai asli atau target yang diketahui. Tujuan dari plotting prediksi adalah untuk memberikan pemahaman visual tentang kualitas dan kecocokan prediksi model dengan data sebenarnya.

3.7 Pengujian Model

Kali ini, uji model, dijalankan eksperimen yang memberi kinerja terbaik pada metode yang dipakai menjadi pemodelan guna menetapkan metode terbaik [18]. Sebelum mencoba metode riset berikut, harus memakai parameter. Lantaran pada analisis dampak indikator berdampak pada keakuratan hasil. Membangun pemodelan perkiraan LSTM sesuai kerangka pembelajaran yang mendalam, yakni platform Python dan Keras Callbacks [27].

3.8 Evaluasi dan Validasi Model

Guna mengevaluasi kinerja model, riset berikut memakai MSE dan RMSE yang dipakai guna pengukuran presisi hasil prediksi model [20]. Validasi dalam metode berikut memakai RMSE dan juga menghasilkan visualisasi yang bertujuan untuk mengetahui tingkat akurasi keseluruhan dari sistem prediksi yang dirancang [28], hingga bisa memecahkan permasalahan yang tampak nyata pada keseharian kehidupan serta berguna dalam memperkirakan di bulan depan. Persamaan yang dipakai guna perhitungan skor MSE dan RMSE ditunjukkan dalam Persamaan 6 dan 7.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{\wedge} - y_i)^2} \quad (6)$$

$$MSE = \frac{1}{n} \sum (y_i - y_i^{\wedge})^2 \quad (7)$$

Dengan keterangan:

1. RMSE = Bobot RMSE
2. y = Bobot hasil pengamatan
3. y^{\wedge} = Bobot hasil perkiraan
4. i = Urutan data dalam database
5. n = Banyaknya data
6. MSE = Rumus MSE