

**PREDIKSI CURAH HUJAN MENGGUNAKAN METODE
*LONG SHORT TERM MEMORY***

(Skripsi)

Oleh

YOGA GUSTAMA



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDARLAMPUNG
2023**

ABSTRAK

PREDIKSI CURAH HUJAN MENGGUNAKAN METODE *LONG SHORT TERM MEMORY*

Oleh

YOGA GUSTAMA

Jaringan Syaraf Tiruan (JST) merupakan sebuah pendekatan pengolahan informasi yang terinspirasi oleh cara kerja sistem syaraf biologis dalam memproses informasi. JST bermanfaat untuk pengenalan pola, *signal processing*, pengklasifikasian dan peramalan Curah hujan merupakan salah satu komponen iklim yang penting dari ekosistem alam. Intensitas iklim bervariasi menurut wilayah. Pada musim hujan, intensitas curah hujan bervariasi, tergantung dari berbagai faktor. Oleh karena itu, memprediksi curah hujan sangat penting untuk berbagai kalangan terutama yang beraktivitas di luar ruangan. Salah satunya dalam melakukan analisis prediksi atau peramalan yang sering digunakan adalah metode *Long Short Term Memory*. LSTM mampu menyimpan informasi pada pola-pola data serta dapat memilih data yang akan digunakan untuk proses selanjutnya dan data yang akan dibuang agar tidak ikut dalam tahapan selanjutnya dalam jaringan. Hal tersebut terjadi karena di dalam *neuron* LSTM terdapat *gate* yang memiliki tiga jenis yakni *input gate*, *forget gate*, dan *output gate* yang dapat mengukur serta memilih memori pada setiap *neuron*. Dari hasil analisis yang telah dilakukan menggunakan algoritma LSTM menghasilkan model terbaik yang dirancang dari 50 hidden neuron, 16 batch size, dan 50 epoch dan nilai RMSE terbaik yaitu 11.43.

Kata kunci: *jaringan syaraf tiruan, curah hujan, long short term memory, prediksi.*

ABSTRACT

RAINFALL PREDICTION USING THE LONG SHORT TERM MEMORY METHOD

By

YOGA GUSTAMA

Artificial Neural Network (ANN) are an information processing approach inspired by the way the biological nervous system works in processing information. ANN is useful for pattern recognition, signal processing, classification and forecasting. Rainfall is an important climate component of natural ecosystems. Climate intensity varies by region. In the rainy season, the intensity of rainfall varies, depending on various factors. Therefore, predicting rainfall is very important for various groups, especially those who do outdoor activities. One of the ways to carry out predictive analysis or forecasting which is often used is the Long Short Term Memory method. LSTM is able to store information on data patterns and can select data that will be used for the next process and data that will be discarded so that it does not participate in the next stage in the network. This happens because in LSTM neurons there are gates which have three types, namely input gate, forget gate and output gate which can measure and sort the memory in each neuron. From the results of the analysis that has been carried out using the LSTM algorithm, it produces the best model designed from 50 hidden neurons, 16 batch sizes, and 50 epochs and the best RMSE value is 11.43.

Keywords: *artificial neural network, rainfall, long short term memory, prediction.*

**PREDIKSI CURAH HUJAN MENGGUNAKAN METODE
*LONG SHORT TERM MEMORY***

Oleh

YOGA GUSTAMA

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2023**

Judul Skripsi : **PREDIKSI CURAH HUJAN MENGGUNAKAN
METODE *LONG SHORT TERM MEMORY***

Nama Mahasiswa : **Yoga Gustama**

Nomor Pokok Mahasiswa : **1917031088**

Jurusan : **Matematika**

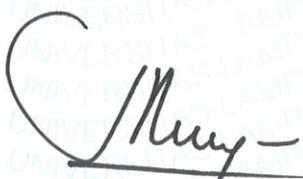
Fakultas : **Matematika dan Ilmu Pengetahuan Alam**




Drs. Nusyirwan, M.Si.
NIP. 196610101992031028


Dra. Dorrah Aziz, M.Si.
NIP. 196101281988112001

2. Ketua Jurusan Matematika


Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316b200501 1 001

MENGESAHKAN

1. Tim Penguji

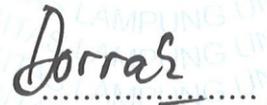
Ketua

: **Drs. Nusyirwan, M.Si.**



Sekretaris

: **Dra. Dorrah Aziz, M.Si.**

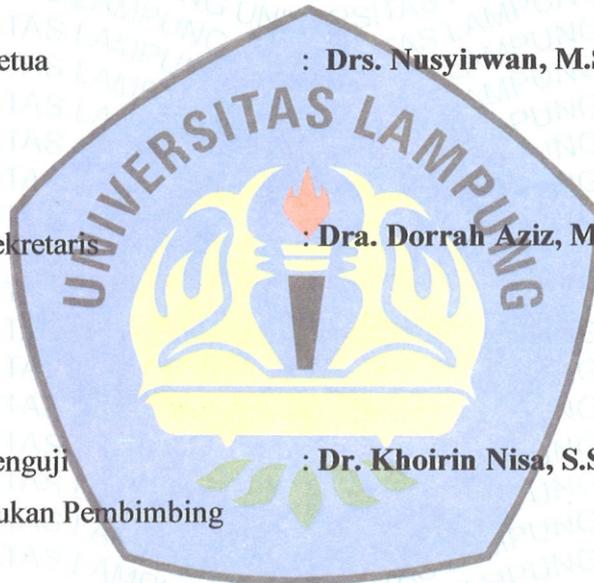


Penguji

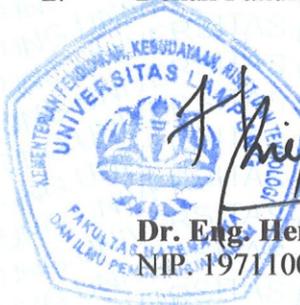
: **Dr. Khoirin Nisa, S.Si., M.Si.**



Bukan Pembimbing



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam




Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : **29 November 2023**

PERNYATAAN SKRIPSI MAHASISWA

Saya bertanda tangan di bawah ini:

Nama : **Yoga Gustama**
Nomor Pokok Mahasiswa : **1917031088**
Jurusan : **Matematika**
Judul : **Prediksi Curah Hujan Menggunakan Metode
*Long Short Term Memory***

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung.

Bandar Lampung, 29 November 2023

Penulis,



Yoga Gustama

RIWAYAT HIDUP

Penulis bernama lengkap Yoga Gustama, anak pertama dari tiga bersaudara yang dilahirkan di Pardasuka pada tanggal 20 September 2000 dari pasangan Bapak Agus Zaniar Roaidi dan Ibu Erma Waty.

Penulis menempuh pendidikan di TK Yustikarini yang diselesaikan pada tahun 2007, kemudian melanjutkan sekolah di SDN 2 Rajabasa yang diselesaikan pada tahun 2013, kemudian melanjutkan sekolah di SMP Negeri 14 Bandar Lampung, dan kemudian melanjutkan sekolah di SMA Negeri 3 Bandar Lampung yang diselesaikan pada tahun 2019.

Pada tahun 2019 penulis terdaftar sebagai Mahasiswa Program Studi S1 Matematika Fakultas Matematika dan Ilmu pengetahuan Alam Universitas Lampung melalui jalur SBMPTN. Selama menjadi mahasiswa penulis pernah aktif menjadi Anggota Bidang Minat dan Bakat Himpunan Matematika (HIMATIKA) 2020.

Pada pertengahan tahun 2022, penulis melaksanakan Kuliah Kerja Nyata (KKN) selama 40 hari di Desa Tanjung Aji, kecamatan Melinting Lampung Timur. Pada awal tahun 2023, penulis melaksanakan Kerja Praktik (KP) di Dinas Ketahanan Pangan, Tanaman Pangan dan Holtikultura Provinsi Lampung.

KATA INSPIRASI

“Janganlah engkau berjalan di bumi ini dengan sombong karena sesungguhnya engkau tidak akan dapat menembus bumi dan tidak akan mampu menjulang setinggi gunung”

(QS. Al-Isra : 37)

“Karena sesungguhnya setelah kesulitan itu ada kemudahan, sesungguhnya setelah kesulitan itu ada kemudahan”

(QS. Al-Insyirah: 5-6)

“Terkadang Tuhan menghancurkan rencanamu sebelum rencanamu menghancurkan dirimu”

(Cak Nun)

“Don't tell people your plans. Show them your results.”

(Jack James)

"Orang yang hebat adalah orang yang memiliki kemampuan menyembunyikan kesusahannya, sehingga orang lain mengira bahwa ia selalu senang."

(Imam Syafi'i)

PERSEMBAHAN

Alhamdulillah Robbil 'Alamin

*Dengan mengucapkan puji dan syukur kepada Allah Subhanahuwata'ala atas
Segala limpah rahmat, taufik dan hidayahNya*

*Kupersembahkan skripsi ini kepada Ayah dan Ibu yang sudah berkorban banyak hal
sehingga aku memiliki kesempatan untuk bahagia bersama kalian. Hanya karena
doa dan didikan kalian yang membawaku bertahan dan kuat sampai sejauh ini.*

*Serta teman-temanku yang memberikan begitu banyak dukungan, doa, canda dan
tawa dalam perkuliahaan. Semua hal yang sudah kita lewati akan kusimpan selalu.*

SANWACANA

Puji syukur kepada Tuhan Yang Maha Esa yang telah memberikan karunia serta kasihNya sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Prediksi Curah Hujan Menggunakan Metode *Long Short Term Memory***”

Dalam proses penyusunan skripsi ini, banyak pihak yang telah membantu memberikan bimbingan, dukungan, motivasi, serta saran sehingga skripsi ini dapat terselesaikan.

Untuk itu penulis ingin mengucapkan terima kasih yang tak terhingga kepada:

1. Ibu Dr. Fitriani, S.Si., M.Si., selaku dosen Pembimbing Akademik yang selalu bersedia memberikan kesediaan waktu, arahan, bimbingan, dan dukungan yang membangun dalam proses penyusunan skripsi ini.
2. Bapak Drs. Nusyirwan, M.Si., selaku Pembimbing I dari awal sampai akhir terimakasih atas kesediaan waktu, saran, dan bimbingan selama proses penyusunan skripsi ini.
3. Ibu Dra. Dorrah Aziz, M.Si., selaku Pembimbing II dari awal sampai akhir terimakasih atas kesediaan waktu, saran, dan bimbingan selama proses penyusunan skripsi ini.
4. Ibu Dr. Khoirin Nisa, S.Si., M.Si. selaku Penguji yang telah bersedia memberikan kritik dan saran serta evaluasi kepada penulis sehingga dapat lebih baik lagi.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Kepala Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh Dosen, Staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Ayah dan Ibu atas cinta, kasih sayang, dukungan dan selalu mendoakan kepada penulis dalam menyelesaikan perkuliahaan.
9. Bintang Gitano Valentino, Surya Anugrah Pratama, Ahmad Yusril Yusro dan Muhammad Deva Randri yang selalu memberikan dukungan dikosan dan doa kepada penulis. Dan seluruh teman-teman jurusan Matematika 2019.
10. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting, just being me at all times.*

Penulis menyadari dalam penyusunan skripsi masih memiliki kekurangan, karena itu kritik dan saran yang membangun diharapkan untuk penyempurnaan skripsi.

Bandar Lampung, 29 November 2023

Penulis,

Yoga Gustama

DAFTAR ISI

	Halaman
DAFTAR ISI	i
DAFTAR TABEL	iii
DAFTAR GAMBAR	iv
I. PENDAHULUAN	1
1.1. Latar Belakang dan Masalah	1
1.2. Tujuan Penelitian.....	3
1.3. Manfaat Penelitian.....	4
II. TINJAUAN PUSTAKA	5
2.1. Peramalan (<i>Forecasting</i>)	5
2.2. <i>Data Cleaning</i>	5
2.3. Normalisasi Data	6
2.4. Pembagian Data	7
2.5. <i>Recurrent Neural Network</i> (RNN)	7
2.6. <i>Long Short Term Memory</i> (LSTM)	8
2.7. Denormalisasi.....	16
2.8. <i>Root Mean Square Error</i> (RMSE)	17
III. METODOLOGI PENELITIAN	18
3.1 Waktu dan Tempat Penelitian	18
3.2 Data Penelitian	18
3.3 Metode Penelitian.....	18
IV. HASIL DAN PEMBAHASAN	21
4.1. Dataset.....	21
4.2. <i>Data Cleaning</i>	22
4.3. Normalisasi.....	24
4.4. Pembagian Data Training dan Data Testing	26
4.5. Membuat Model LSTM	26
4.6. Perhitungan Manual	27
4.7. Validasi Model	31
4.8. Prediksi.....	36

V. KESIMPULAN DAN SARAN 39

5.1. Kesimpulan.....39

5.2. Saran.....39

DAFTAR PUSTAKA.....40**LAMPIRAN**

DAFTAR TABEL

Tabel	Halaman
1. Dataset curah hujan.....	22
2. Jumlah <i>missing value</i> pada data curah hujan	23
3. Data curah hujan setelah dilakukan data <i>cleaning</i>	23
4. Nilai minimum dan maksimum data curah hujan	24
5. Hasil normalisasi data curah hujan	25
6. Pembagian data <i>training</i> dan data <i>testing</i>	26
7. Nilai <i>input</i> , <i>weight</i> , dan bias pada <i>timestep</i> pertama.....	27
8. Nilai akurasi model LSTM berdasarkan jumlah <i>hidden neuron</i>	31
9. Nilai akurasi model LSTM berdasarkan jumlah <i>batch size</i>	32
10. Nilai akurasi model LSTM berdasarkan jumlah <i>epoch</i>	33
11. Nilai akurasi model LSTM berdasarkan pembagian <i>training</i> dan <i>testing</i>	33
12. Model LSTM curah hujan terbaik	34
13. Hasil peramalan curah hujan	37

DAFTAR GAMBAR

Gambar	Halaman
1. Arsitektur <i>Recurrent Neural Network</i>	8
2. Perulangan pada RNN dengan satu layer.....	9
3. Struktur jaringan (LSTM).....	10
4. <i>Cell state</i> pada LSTM.....	11
5. Alur <i>forget gate</i> pada LSTM.....	12
6. Alur <i>input gate</i> pada LSTM.....	13
7. Alur memperbarui <i>cell state</i> pada LSTM.....	14
8. Alur <i>output</i> pada LSTM.....	15
9. Diagram alir metode penelitian.....	20
10. Plot data curah hujan.....	21
11. Plot data <i>cleaning</i> curah hujan.....	24
12. Plot prediksi curah hujan.....	36
13. Plot peramalan curah hujan.....	37

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Artificial Neural Network (ANN) atau Jaringan Syaraf Tiruan (JST) merupakan sebuah pendekatan pengolahan informasi yang terinspirasi oleh cara kerja sistem syaraf biologis dalam memproses informasi. ANN bermanfaat untuk pengenalan pola, *signal processing*, pengklasifikasian dan peramalan (Siang, 2009). Jaringan syaraf tiruan (ANN) terbentuk dari sekelompok *neuron* yang terstruktur dalam beberapa tingkatan, termasuk lapisan *input* (*input layer*), lapisan tersembunyi (*hidden layer*), dan lapisan *output* (*output layer*). Dalam penerapannya, penggunaan jaringan syaraf tiruan (ANN) dalam menganalisis seringkali dilakukan untuk memprediksi atau meramalkan data sekuensial, terutama dalam konteks data runtun waktu. Salah satu algoritma yang umum digunakan untuk analisis prediksi atau peramalan adalah *Recurrent Neural Network*.

RNN merupakan suatu jenis arsitektur jaringan syaraf tiruan yang beroperasi dengan cara memproses *input* secara berulang. Data yang umumnya dimanfaatkan dalam struktur RNN ini adalah data sekuensial. Data sekuensial memiliki ciri khas di mana setiap sampel diproses secara berurutan, seperti berdasarkan waktu, dan setiap sampel dalam urutan tersebut memiliki keterkaitan yang kuat satu sama lain. RNN sesuai untuk digunakan dalam mempelajari pola suatu data. *Long Short Term Memory* (LSTM) merupakan salah satu varian atau jenis RNN. LSTM mampu menyimpan informasi pada pola-pola data serta dapat memilih data yang akan digunakan untuk proses selanjutnya dan data yang akan dibuang agar tidak ikut dalam tahapan

selanjutnya dalam jaringan. Hal tersebut terjadi karena di dalam *neuron* LSTM terdapat *gate* yang memiliki tiga jenis yakni *input gate*, *forget gate*, dan *output gate* yang dapat mengukur serta memilah memori pada setiap *neuron*. Jaringan LSTM ini banyak digunakan dalam pemrosesan data *time series*, pemrosesan teks, video, dan lain-lain (Aldi, dkk., 2018).

Peramalan merupakan teknik untuk meramalkan atau memprediksi keadaan di masa yang akan datang melalui pengujian keadaan di masa lalu. Pada dasarnya meramalkan sama halnya dengan memprediksi atau memperkirakan sesuatu hal, peristiwa masa datang yang berdasarkan data pada masa lalu (runtun waktu) hingga saat ini. Dalam melakukan peramalan analisis deret waktu dibutuhkan model statistik yang tepat untuk menjelaskan suatu data deret waktu. Hal ini tentu tepat digunakan pada data curah hujan di Indonesia yang merupakan data deret waktu yang bisa diramalkan ke depannya.

Curah hujan merupakan salah satu komponen iklim yang penting dari ekosistem alam. Garis khatulistiwa yang terbentang di Indonesia menjadikan Indonesia sebagai negara yang beriklim tropis. Musim kemarau dan musim hujan adalah dua musim di lingkungan tropis. Intensitas iklim bervariasi menurut wilayah. Misalnya, pada musim hujan, intensitas curah hujan bervariasi tergantung pada garis lintang wilayah, ketinggian, kedekatan dengan sumber air, arah angin, suhu tanah, dan durasi curah hujan. Karena adanya berbagai unsur yang dapat mempengaruhi besarnya curah hujan, maka cuaca di Indonesia tidak selalu berjalan normal atau sesuai dengan musim. Sebaliknya, itu sering berubah tiba-tiba setiap saat.

Curah hujan di Indonesia memiliki rata-rata yang mencerminkan keadaan curah hujan normal. Alat yang digunakan untuk mengukur curah hujan melibatkan perangkat pengukur curah hujan, radar berbasis darat, dan satelit penginderaan jauh. Pola hujan di Jawa bersifat musiman, dengan karakteristik unimodal, yakni memiliki puncak musim hujan. Banyak studi kasus yang terkait peramalan curah hujan yang telah dilakukan dengan berbagai metode, termasuk penggunaan Decision Tree (Hasanah,

dkk., 2021), Model Rantai Markov (Ihsan, dkk., 2019), dan Persamaan Regresi (Putramulyo & Alaa, 2018).

Beberapa penelitian telah banyak menggunakan metode *Long Short Term Memory* diantaranya adalah prediksi curah hujan Kota Malang (Rizki, dkk., 2020), analisis peramalan kasus Covid-19 di Jawa Tengah (Qori, dkk., 2022), prediksi Harga Bitcoin (Aldi, dkk., 2018), dan sebagainya.

Penelitian mengenai pendekatan LSTM ini merujuk pada studi yang dilakukan oleh Rizki dan rekan-rekannya pada tahun 2020, yang fokus pada penerapan *Deep Learning* dengan menggunakan arsitektur LSTM untuk meramalkan curah hujan di Kota Malang. Dalam penelitian tersebut, perbandingan dilakukan terhadap model LSTM dengan variasi parameter LSTM, seperti *hidden neuron*, *batch*, dan *epoch*. Hasil menunjukkan bahwa model LSTM yang paling optimal dalam memprediksi curah hujan di Kota Malang memiliki nilai RMSE terendah, yakni pada model dengan *hidden neuron* sebanyak 256 dan *epoch* sebanyak 150 (Rizki, dkk., 2020).

Metode peramalan secara umum banyak memerlukan pemenuhan beberapa asumsi tertentu. Namun, ada juga jenis model yang tidak bergantung pada asumsi-asumsi tersebut, dan salah satu contohnya adalah Jaringan Syaraf (*Neural Network*) (Rinjani, dkk., 2019). Oleh karena itu, penulis berniat untuk menganalisis data deret waktu dengan menggunakan metode *Long Short Term Memory* (LSTM) guna melakukan prediksi curah hujan.

1.2 Tujuan Penelitian

Adapun tujuan penulisan skripsi adalah sebagai berikut:

1. Mengetahui struktur jaringan yang optimal yang dihasilkan oleh metode *Long Short Term Memory* pada data curah hujan.

2. Mengetahui hasil peramalan curah hujan berdasarkan data *time series* curah hujan yang dihasilkan dari model LSTM yang terbaik.

1.3 Manfaat Penelitian

Diharapkan penulisan penelitian ini dapat berfungsi sebagai sumber referensi untuk penerapan metode *Long Short Term Memory* dalam peramalan, menjadi acuan dalam penelitian selanjutnya, dan memberikan informasi yang bermanfaat. Selain itu, diharapkan penelitian ini dapat menjadi pertimbangan penting dalam melakukan analisis peramalan curah hujan.

II. TINJAUAN PUSTAKA

2.1 Peramalan

Menurut Qori (2022), peramalan, atau sering disebut sebagai *forecasting*, adalah ilmu yang mempelajari estimasi suatu peristiwa berdasarkan data-data masa lalu. Analisis peramalan dilakukan dengan merujuk pada data historis, khususnya data deret waktu (*time series*), sehingga memungkinkan untuk memprediksi kejadian di masa depan. Metode peramalan mencakup dua pendekatan utama, yakni pendekatan kualitatif dan pendekatan kuantitatif.

Pendekatan kualitatif adalah suatu metode peramalan yang bergantung pada pendapat atau opini pihak terkait, terutama saat tidak ada informasi historis yang dapat digunakan sebagai dasar estimasi. Sementara itu, metode peramalan dengan pendekatan kuantitatif adalah suatu pendekatan yang memanfaatkan data historis sebagai dasar estimasi, seperti yang dijelaskan oleh Montgomery (2008).

2.2 Data Cleaning

Teknik pembersihan data, atau yang dikenal sebagai data *cleaning*, digunakan untuk mengatasi masalah nilai yang hilang atau *noise* dalam suatu dataset. Keadaan nilai yang hilang, atau yang disebut sebagai *missing value*, terjadi ketika suatu dataset tidak memiliki data atau informasi yang lengkap. Terdapat dua metode dalam data *cleaning* untuk menangani *missing value* (Aprian, dkk., 2020):

1. Menghapus seluruh baris (*row*) pada data yang memiliki nilai yang hilang
2. Menggantikan nilai yang hilang dengan nilai konstan, rata-rata, *median*, atau modus untuk kolom yang mengalami kekurangan data, atau melakukan beberapa metode imputasi lainnya.

Rumus untuk menghitung nilai rata-rata (*mean*) dari suatu dataset dapat dinyatakan dalam persamaan berikut:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) \quad (2.2)$$

2.3 Normalisasi Data

Normalisasi data atau *scaling* data adalah prosedur yang mengubah format data asli menjadi bentuk lain. Normalisasi data adalah langkah penting dalam upaya mencapai nilai *error* yang seminim mungkin. Dalam konteks peramalan curah hujan, teknik normalisasi yang diterapkan adalah *Min Max Scaler*, yang juga dikenal sebagai *Min-Max Scaling*. Teknik ini memodifikasi data aktual sehingga nilainya terdistribusi dalam rentang [0,1].. Teknik *Min-Max Scaling* dirumuskan ke dalam persamaan (2,3) berikut (Rizki, dkk., 2020):

$$x' = \frac{x - X_{min}}{X_{max} - X_{min}} \quad (2.3)$$

Keterangan:

x' : Data curah hujan yang telah dinormalisasikan.

x : Data aktual curah hujan.

X_{min} : Nilai minimum data curah hujan.

X_{max} : Nilai maksimum data curah hujan.

2.4 Pembagian Data

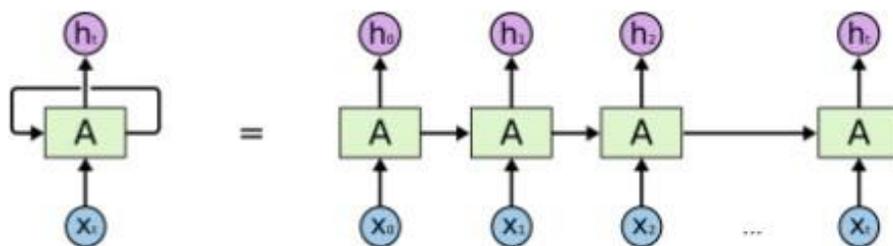
Dalam pengembangan model *Long Short Term Memory* (LSTM), data merupakan elemen kunci. Kumpulan data yang digunakan dalam jaringan syaraf tiruan disebut sebagai dataset, yang selanjutnya dibagi menjadi dua bagian: *training* dataset dan *testing* dataset. Proporsi pembagian data dapat disesuaikan untuk mencari pengaturan terbaik, seperti 50:50, 60:40, 70:30, dan 80:20, dengan tujuan mencapai pembagian data yang optimal (Utami & Ulama, 2015).

2.5 Recurrent Neural Network

Recurrent Neural Network (RNN) merupakan jenis *Artificial Neural Network* yang dirancang khusus untuk memproses data sekuensial. RNN memiliki kemampuan untuk mengolah informasi dari masa lalu dalam proses pembelajarannya. Karena kemampuannya dalam memproses data sekuensial, RNN dapat diterapkan dalam penanganan data *time series*. Struktur RNN terdiri dari *input layer*, *hidden layer*, dan *output layer*. Aliran informasi dalam RNN bersifat satu arah, dimulai dari *input layer* ke *hidden layer*, dan informasi mengalir dari *hidden layer* sebelumnya ke *hidden layer* saat ini. Dengan demikian, *output* dari *hidden layer* menjadi *input* untuk langkah berikutnya (Larasati & Primandari, 2021).

Dalam proses memprediksi, RNN memanfaatkan data *input* pada waktu sekarang bersama dengan *input* dari periode sebelumnya. Hubungan antara kedua jenis *input* ini memberikan kontribusi informasi kepada seluruh lapisan tersembunyi (*hidden layer*). Oleh karena itu, dapat diungkapkan bahwa RNN memiliki kemampuan memori, yang menyimpan rekaman hasil informasi yang dihasilkan pada langkah-langkah sebelumnya (Gulli, 2017).

Meskipun secara konseptual dan berdasarkan struktur teoritis RNN dianggap dapat menangani ketergantungan jangka panjang, dalam prakteknya RNN sering kali mengalami kesulitan dalam menangani ketergantungan jangka panjang. Hal ini terjadi karena RNN menghadapi masalah *gradien* yang hilang, yang menyebabkan ketidakmampuannya menyimpan informasi dari masa lalu dengan baik. Untuk mengatasi masalah ini, Hochreiter & Schmidhuber pada tahun 1997 memperkenalkan varian khusus dari RNN yang dikenal sebagai *Long Short-Term Memory* (LSTM). LSTM dirancang agar dapat mengatasi tantangan ketergantungan jangka panjang yang dihadapi oleh RNN konvensional (Larasati & Primandari, 2021).

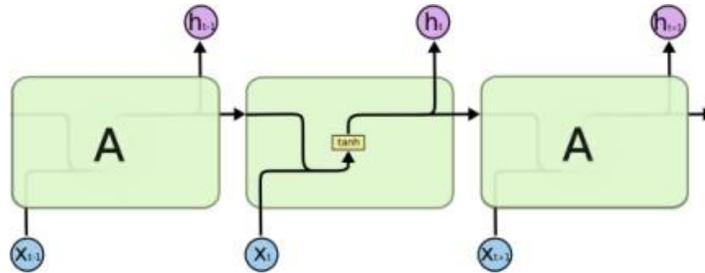


Gambar 1. Arsitektur *Recurrent Neural Network* (Colah, 2015)..

2.6 *Long Short Term Memory*

Long Short Term Memory (LSTM) merupakan *evolusi* dari arsitektur *Recurrent Neural Network* (RNN). Hochreiter & Schmidhuber adalah inovator di balik pengembangan metode LSTM, yang pertama kali diperkenalkan kepada publik pada tahun 1997. Kemunculan LSTM dipicu oleh ketidakpuasan terhadap kemampuan arsitektur RNN dalam memproses data sekuensial dengan durasi panjang. Kekurangan utama RNN adalah masalah *gradien* yang hilang, terutama saat menerapkan algoritma *backpropagation* (Wang, 2022).

LSTM dianggap sebagai perkembangan dari RNN karena pada dasarnya keduanya memiliki struktur serupa, melibatkan *input layer*, *hidden layer*, dan *output layer*. Perbedaan utama terletak pada konfigurasi jaringan di lapisan tersembunyi (*hidden layer*). Pada RNN, *hidden layer* terdiri dari satu lapisan sederhana dengan menggunakan fungsi aktivasi Tanh seperti Gambar 2 berikut :



Gambar 2. Perulangan pada RNN dengan satu *layer* (Colah, 2015).

Pada Gambar 2 di atas adalah perulangan pada RNN yang hanya terdiri dari satu *layer* yakni *tanh layer* yang akan dijabarkan pada persamaan (2.4) berikut :

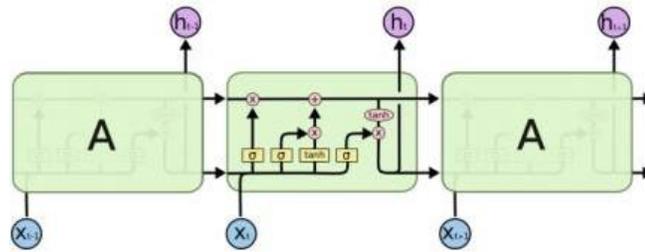
$$\text{Tanh}(x) = 2\sigma(2x) - 1 \quad (2.4)$$

Dimana,

σ = Fungsi Sigmoid

x = Data *input*

Sementara itu, dalam struktur LSTM terdapat perbedaan karena terdiri dari dua fungsi aktivasi, yaitu fungsi Sigmoid dan Tanh. Lebih lanjut, LSTM melibatkan komponen-komponen seperti *memory cell* dan *gates*, di mana *gates* tersebut terdiri dari tiga bagian, yakni *forget gate*, *input gate*, dan *output gate*. Ilustrasi di bawah menjelaskan secara *visual* struktur jaringan LSTM.

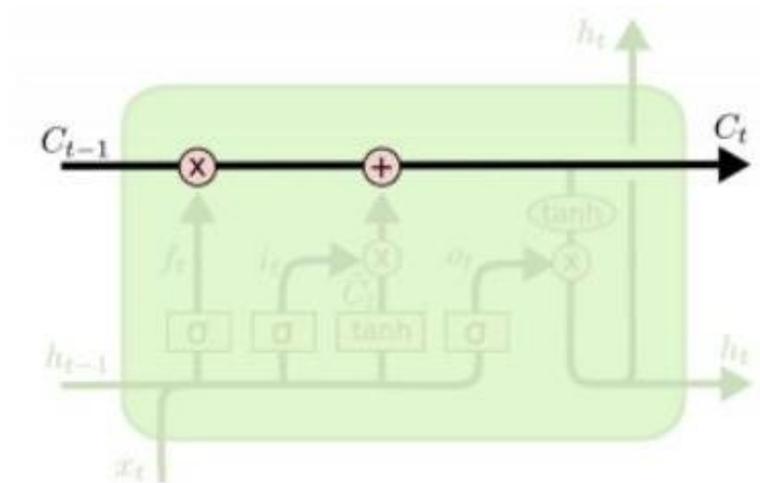


Gambar 3. Struktur jaringan LSTM (Colah, 2015).

Persamaan metode LSTM pada Gambar 3 diuraikan menjadi persamaan (2.5), sebagai berikut.

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f) \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\
 \bar{C}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c) \\
 C_t &= f_t * C_{t-1} + i_t * \bar{C}_t \\
 o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\
 h_t &= o_t * \tanh(C_t)
 \end{aligned}
 \tag{2.5}$$

Langkah-langkah yang terlihat pada Gambar 3 akan dijelaskan secara detail pada bagian ini. Komponen kunci dalam metode LSTM adalah *cell state*. *Cell state* merupakan garis *horizontal* yang menyilang di bagian atas diagram, menghubungkan seluruh *output layer* pada LSTM, seperti yang tergambar pada Gambar 4.



Gambar 4. *Cell state* pada LSTM (Colah, 2015)

LSTM dilengkapi dengan *gates* yang terdiri dari lapisan sigmoid dan operasi *pointwise*. *Gates* ini memiliki kemampuan untuk menambah atau menghapus informasi, dan apakah informasi tersebut akan diteruskan atau dihentikan. Sigmoid *layer* berfungsi untuk memperkecil hasil *output* ke dalam rentang $[0,1]$, dengan nilai 0 menandakan bahwa informasi akan dihentikan, sementara nilai 1 menandakan bahwa informasi akan diteruskan. Persamaan Sigmoid dapat dijabarkan menggunakan persamaan (2.6) berikut.

$$\sigma(x) = \frac{1}{(1 + e^{-x})} \quad (2.6)$$

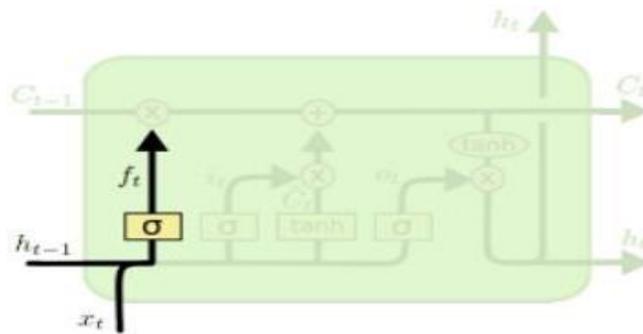
Dimana,

x = Data *input*

e = Konstanta matematika (2,71828 18284 59045 23536 02874 71352)

Proses langkah dalam metode LSTM adalah sebagai berikut:

Langkah awal terfokus pada *forget gate*. *Forget gate* ini bertanggung jawab untuk menentukan informasi yang akan dihapus dari *cell state*. Proses ini terjadi melalui lapisan sigmoid, yang menghasilkan *output* berupa angka antara 0 dan 1. Perhitungan *forget gate* dilakukan dengan memanfaatkan *output* data sebelumnya (h_{t-1}) dan data *input* saat ini (x_t) seperti yang terlihat pada Gambar 5.



Gambar 5. Alur *forget gate* pada LSTM (Colah, 2015)

Persamaan *forget gate* diuraikan pada persamaan (2.7) berikut.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.7)$$

Dimana,

f_t = *Forget gate*

σ = Fungsi Sigmoid

W_f = Nilai *weight* untuk *forget gate*

h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_f = Nilai bias pada *forget gate*

Nilai *weight* dirumuskan dengan persamaan (2.8).

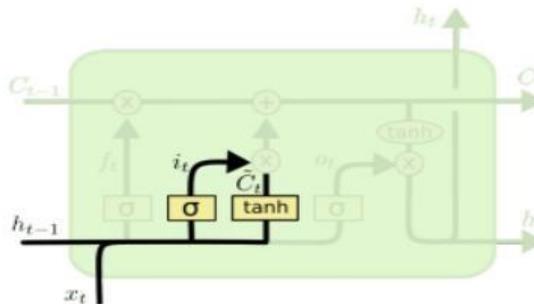
$$W = \left(-\frac{1}{\sqrt{d}}, \frac{1}{\sqrt{d}}\right) \quad (2.8)$$

Dimana,

W = *Weight*

d = Jumlah variabel

Tahap ini diilustrasikan pada Gambar 6



Gambar 6. Alur *input gate* pada LSTM (Colah, 2015)

Tahapan kedua *input gate* yang merupakan proses mengolah informasi untuk menentukan informasi yang diperbarui dan kemudian dicadangkan ke dalam status cell.

Persamaan pada *input gate* ditulis seperti persamaan (2.9) berikut.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.9)$$

Dimana,

i_t = *Input gate*

σ = Fungsi Sigmoid

W_i = Nilai *weight* untuk *input gate*

h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_i = Nilai bias pada *input gate*

Persamaan kandidat baru ditulis pada persamaan (2.10).

$$\bar{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.10)$$

Dimana:

\bar{C}_t = Nilai baru yang dapat ditambahkan ke *cell state*

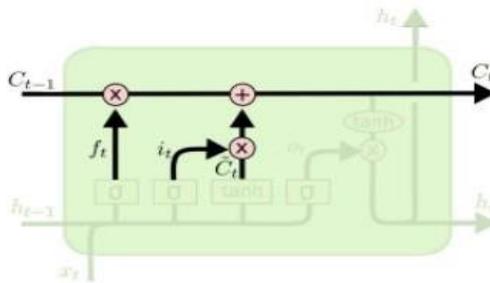
\tanh = Fungsi tangen hiperbolik

W_c = Nilai *weight* untuk *cell state*

h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_c = Nilai bias pada *cell state*



Gambar 7. Alur memperbarui *cell state* pada LSTM (Colah, 2015)

Langkah ketiga melibatkan pembaruan *cell state*, yang dijelaskan dalam Gambar 7. Dalam ilustrasi tersebut, pembaruan *cell state* terjadi dengan mengalikan *output* dari *forget gate* dengan *cell state* sebelumnya, dan hasilnya ditambahkan dengan hasil dari langkah kedua, yakni $i_t * \bar{C}_t$. Persamaan *cell state* tersebut ditampilkan pada persamaan (2.11)

$$C_t = f_{t-1} * C_{t-1} + i_t * \bar{C}_t \quad (2.11)$$

Dimana:

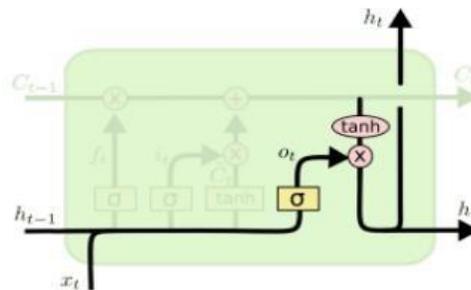
C_t = Cell state

f_t = Forget gate

C_{t-1} = Cell state sebelum orde ke t

i_t = Input gate

\bar{C}_t = Nilai baru yang dapat ditambahkan ke cell state



Gambar 8. Alur *output* pada LSTM (Colah, 2015).

Langkah terakhir dalam metode LSTM ini adalah menentukan hasil *output*, sebagaimana diilustrasikan dalam Gambar 8. Lapisan sigmoid menentukan bagian dari *cell state* yang akan dijadikan *output*. Persamaan (2.12) menggambarkan operasi *output gate* dalam proses ini.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.12)$$

Dimana,

o_t = Output gate

σ = Fungsi Sigmoid

w_o = Nilai *weight* untuk *output gate*

h_{t-1} = Nilai *output* sebelum orde ke t

x_t = Nilai *input* pada orde ke t

b_o = Nilai bias pada *output gate*

Setelah mendapatkan nilai dari *output gate*, langkah selanjutnya adalah memasukkan *cell state* melalui fungsi tanh untuk mendapatkan nilai dalam rentang -1 hingga 1. Kemudian, nilai ini dikalikan dengan *output gate* dari lapisan sigmoid. Persamaan untuk nilai *output* pada orde t dapat diungkapkan dalam persamaan (2.13).

$$h_t = o_t * \tanh(C_t) \quad (2.13)$$

Dimana:

h_t = Nilai *output* orde t

o_t = *Output gate*

tanh = Fungsi tangen hiperbolik

C_t = *Cell state*

2.7 Denormalisasi

Denormalisasi adalah langkah untuk mengembalikan data yang telah dinormalisasi ke dalam skala data asli atau skala sebenarnya. Tujuan dari proses ini adalah untuk memungkinkan perbandingan hasil prediksi dengan data aktual (Wiranda & Sadikin, 2019). Denormalisasi dalam rentang [0, 1] dapat dihitung menggunakan rumus yang tertera dalam persamaan (2.14) sebagai berikut:

$$x_t = x'(X_{max} - X_{min}) + X_{min} \quad (2.14)$$

Dimana:

X_t = Nilai dari data normalisasi

x' = Hasil output

X_{min} = Nilai minimal data aktual keseluruhan

X_{max} = Nilai maksimal data aktual keseluruhan

2.8 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) adalah metode evaluasi untuk menilai tingkat akurasi model regresi linear dengan mengukur seberapa dekat prediksi model dengan observasi aktual. Perhitungan RMSE melibatkan langkah-langkah seperti mengkuadratkan selisih antara prediksi dan *observasi*, kemudian hasilnya dijumlahkan dan dibagi dengan jumlah data (rata-rata), dan akhirnya diambil akar kuadrat. Penggunaan RMSE dapat menghindari kelemahan nilai absolut yang tidak diinginkan dalam berbagai perhitungan matematis. Persamaan untuk RMSE dapat dijelaskan melalui persamaan (2.15) berikut.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \quad (2.15)$$

Dimana:

Y_i = Data aktual pada periode i

\hat{Y}_i = Data prediksi pada periode i

n = Jumlah periode peramalan yang terlibat

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2022/2023 bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

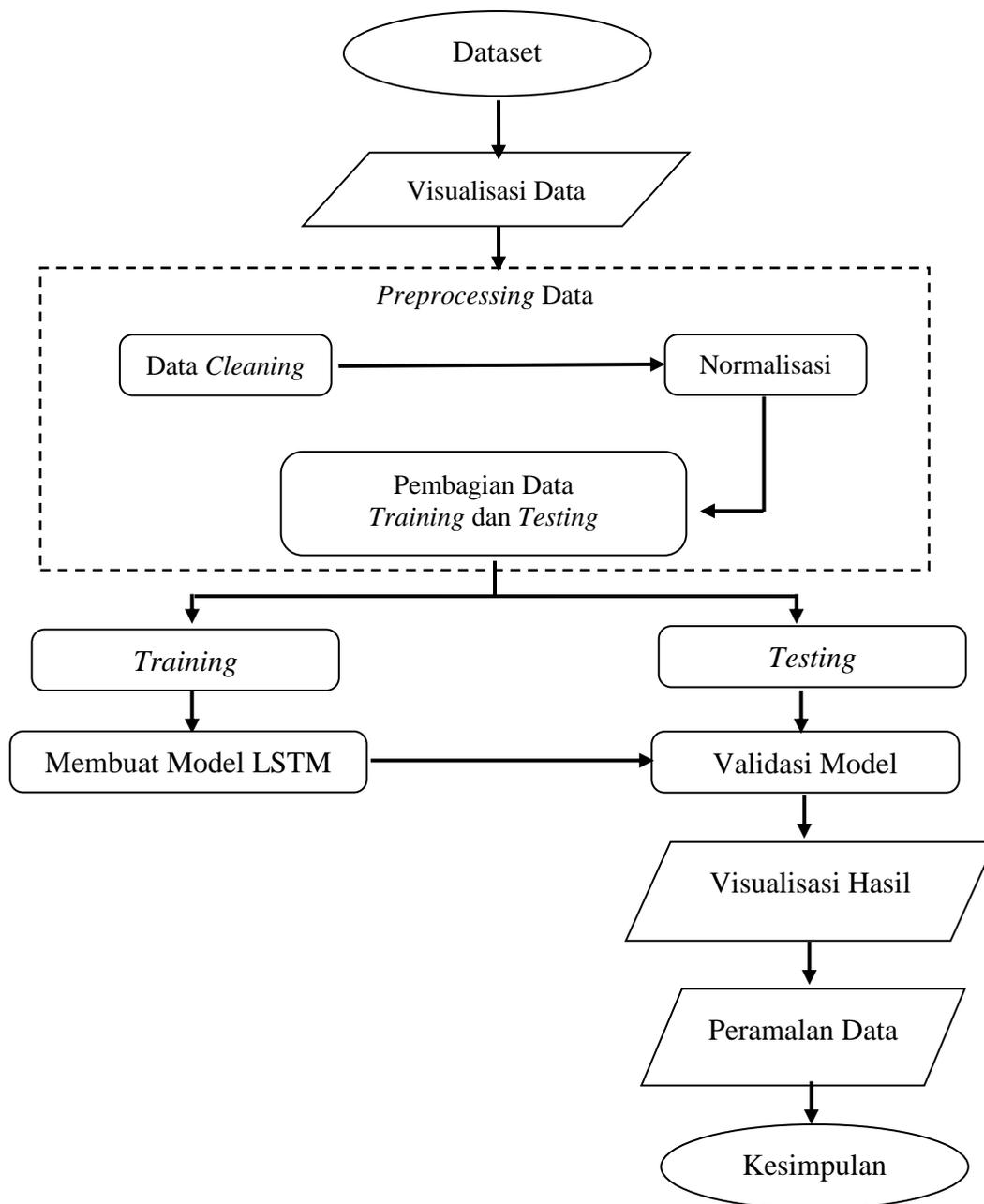
3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data curah hujan Kabupaten Bogor periode 1 Juli 2021 sampai dengan 30 Juli 2023 yang diperoleh dari <https://dataonline.bmkg.go.id> pada Stasiun Meteorologi Ciketo dengan jumlah data sebanyak 760 data. Dalam studi ini, dataset terbagi menjadi dua bagian, yakni data *training* yang digunakan untuk melatih model, dan data *testing* yang digunakan untuk mengevaluasi model yang terbentuk. Pembagian dataset dilakukan dalam empat bagian untuk memastikan proses yang baik dalam pembentukan dan pemilihan model terbaik.

3.3 Metode Penelitian

Dalam penelitian ini dilakukan analisis menggunakan metode *long short term memory*. Langkah-langkah yang dilakukan yaitu:

1. Input data historis curah hujan.
2. Visualisasi data dengan grafik *time series*.
3. Melakukan *preprocessing* data yang mencakup data *cleaning* untuk menangani nilai yang hilang atau *missing value*, normalisasi data menggunakan pendekatan Min-Max Scaling, dan membagi dataset menjadi dua bagian, yaitu data untuk pelatihan (*training*) dan data untuk pengujian (*testing*).
4. Pembentukan model LSTM dilakukan dengan menggunakan data *training*..
5. Melakukan validasi model menggunakan data *testing* untuk mendapatkan nilai RMSE.
6. Memeriksa performa model yang dibangun dengan variasi jumlah *hidden neuron* (5, 25, 50, 100), *batch Size* (4, 16, 32, 64), dan jumlah *epoch* (50, 100, 150, 200). Model LSTM yang memiliki nilai RMSE terendah akan dianggap sebagai model terbaik.
7. Menampilkan visualisasi hasil prediksi curah hujan berdasarkan hasil model terbaik.
8. Melakukan prediksi untuk beberapa waktu mendatang berdasarkan hasil dari model terbaik.
9. Kesimpulan.



Gambar 9. Diagram alir metode penelitian

V. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Setelah menganalisis, merancang, dan mengimplementasikan serta menguji hasil dengan menggunakan struktur *Long Short Term Memory* (LSTM) untuk meramalkan curah hujan harian di Kabupaten Bogor, diperoleh kesimpulan dari penelitian ini, diantaranya:

1. Model LSTM yang optimal untuk meramalkan curah hujan di Kabupaten Bogor didasarkan pada pengujian pembagian dataset dengan komposisi 80% untuk data training dan 20% untuk data testing.
2. Model LSTM terunggul untuk meramalkan curah hujan di Kabupaten Bogor dirancang dari 50 *hidden neuron*, 16 *batch size*, dan 50 *epoch*.
3. RMSE dari model LSTM yang optimal untuk meramalkan curah hujan di Kabupaten Bogor adalah sebesar 11,43.

5.2. Saran

Peneliti berharap bahwa penelitian ini dapat menjadi landasan untuk pengembangan lebih lanjut oleh peneliti lain di masa depan. Adapun saran yang diberikan oleh peneliti adalah

1. Menambahkan beberapa parameter atau metode tambahan untuk meningkatkan kualitas penelitian dan memungkinkan perbandingan lebih mendalam untuk menentukan metode yang paling efektif.

DAFTAR PUSTAKA

- Aldi, M. W. P., Jondri., & Annisa, A. 2018. Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin, e-Proceeding of Engineering, **5(2)**: 3548-3555.
- Aprian, B. A., Azhar, Y. & Nastiti, V. R. S. 2020. Prediksi Pendapatan Kargo Menggunakan Arsitektur Long Short Term Memory. *Jurnal Komputer Terapan*. **6(2)**: 148-157.
- Badriyah, J., Fariza, A., & Harsono, T. (2022). Prediksi Curah Hujan Menggunakan Long Short Term Memory. *Jurnal Media Informatika Budidarma*. **6(3)**: 1297-1303.
- Colah. 2015. "Understanding LSTM Networks". <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Diakses pada 22 Januari 2021 jam 12.32 WITA.
- Gulli. A dan S. P. 2017. *Deep Learning with Keras*, Birmingham: Packt Publishing Ltd.
- Hasanah, M. A., Soim, S., Handayani, A.S. 2021. Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. *Journal of Applied Informatics and Computing (JAIC)*. **5(2)**: 103-108.
- Hayuningtyas, R. Y. 2017. Peramalan Persediaan Barang Menggunakan Metode Weighted Moving Average dan Metode Double Exponential Smoothing. *Jurnal PILAR Nusa Mandiri*. **13(2)**: 217-219.

- Ihsan, H., Sanusi, W., dan Hasriani. 2019. Peramalan Pola Curah Hujan Di Kota Makassar Menggunakan Model Rantai Markov. *Journal of Mathematics, Computations, and Statistics*. 2(1): 19-30.
- Larasati, K. D. & Primandari, A. H. 2021. Forecasting Bitcoin Price Based on Blockchain Information Using Long-Short Term Method. *Journal of Statistics*. 1(1).
- Montgomery, D. C., Jennings, L. & Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. John Wiley And Sons Inc, Canada.
- Putramulyo, S. & Alaa, S. 2018. Prediksi Curah Hujan Bulanan Di Kota Samarinda Menggunakan Persamaan Regresi Dengan Prediktor Data Suhu dan Kelembapan Udara. *Eigen Mathematics Journal*. 1(2).
- Qori, P. A., Oktavani, D. S. & Kharisudin, I. 2022. Analisis Peramalan dengan Long Short Term Memory pada Data Kasus Covid-19 di Provinsi Jawa Tengah. PRISMA, Prosiding Seminar Nasional Matematika. 5: 752-758.
- Rinjani, S.N., Hoyyi, A. & Suparti. 2019. Pemodelan Fungsi Transfer dan Backpropagation Neural Network untuk Peramalan Harga Emas. *Jurnal Gaussian*. 8(4): 474-485.
- Rizki, M., Basuki, S. & Azhar, Y. 2020. Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory untuk Prediksi Curah Hujan Kota Malang. *Jurnal Repositor*. 2(3): 331-338.
- Siang, Jong Jek, (2009). *Jaringan Syaraf Tiruan dan Pemrograman Menggunakan MATLAB*, Yogyakarta: ANDI.
- Utami, A. T. W. & Ulama, B. S. S. 2015. Penerapan Backpropagation untuk Meningkatkan Efektivitas Waktu dan Akurasi pada Data Wall-Following Robot Navigation. *Jurnal Sains dan Seni ITS*. 4(2): 279-284.

Wang, J., Li, X., Li, J., Sun, Q., & Wang, H. (2022). NGCU: A New RNN Model for Time-Series Data Prediction. *Big Data Research*. **27**.

Wiranda, L. & Sadikin, M. 2019. Penerapan Long Short Term Memory pada Data Time Series untuk Memprediksi Penjualan Produk PT. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika*. **8**(3): 184-196.

Zheng J and Huang M. 2020. Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning. *IEEE Access*. **8**: 82562-82570.