

**PEMODELAN *TIME SERIES* UNTUK PERAMALAN SUHU UDARA
MENGUNAKAN METODE *LONG SHORT TERM MEMORY* (LSTM)
(Studi Kasus: Stasiun Klimatologi Lampung)**

(Skripsi)

Oleh

**LINDA SUNDARI
1917031070**



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

ABSTRACT

TIME SERIES MODELING FOR AIR TEMPERATURE FORECASTING USING LONG SHORT TERM MEMORY (LSTM) METHOD (Case Study: Klimatologi Lampung Station)

By

LINDA SUNDARI

Forecasting is a science that studies an event in the future based on historical data in the form of time series data. Time series data is data taken from time to time (sequential data). Forecasting analysis on time series data can be performed using the Long Short Term Memory (LSTM) method. This method is capable of storing input data information for a long time based on past data information. This study aims to model time series data and determine the performance of LSTM in forecasting air temperature. Testing the accuracy of this model uses RMSE and MAPE values, where the model that has the smallest RMSE and MAPE values is the best model. The best LSTM model in air temperature forecasting uses 80% training data and 20% testing data. The LSTM model for forecasting Minimum Air Temperature is formed from 50 hidden neurons, 4 batch sizes, and 50 epochs with an RMSE value of 0.89 and a MAPE of 2.59%. While the best LSTM model for forecasting Maximum Air Temperature and Average Air Temperature is formed from 25 hidden neurons, 4 batch sizes, and 50 epochs with maximum RMSE and MAPE Air Temperature values of 0.85 and 2.32% and Average Air Temperature average of 0.88 and 2.44%. So this LSTM model is very good for air temperature forecasting.

Keywords: Forecasting, Time Series, Long Short Term Memory (LSTM).

ABSTRAK

PEMODELAN *TIME SERIES* UNTUK PERAMALAN SUHU UDARA MENGUNAKAN METODE *LONG SHORT TERM MEMORY* (LSTM) (Studi Kasus: Stasiun Klimatologi Lampung)

Oleh

LINDA SUNDARI

Peramalan merupakan ilmu yang mempelajari suatu peristiwa pada masa yang akan datang berdasarkan data historis yang berupa data *time series*. Data *time series* adalah data yang diambil berdasarkan urutan waktu ke waktu (data sekuensial). Analisis peramalan pada data *time series* dapat dilakukan dengan menggunakan metode *Long Short Term Memory* (LSTM). Metode ini mampu menyimpan informasi data masukan dalam waktu yang panjang berdasarkan informasi data pada masa lalu. Pada penelitian ini bertujuan untuk memodelkan data *time series* dan mengetahui performa LSTM dalam melakukan peramalan suhu udara. Pengujian akurasi terhadap model ini menggunakan nilai RMSE dan MAPE, dimana model yang memiliki nilai RMSE dan MAPE yang terkecil merupakan model yang terbaik. Model LSTM yang terbaik dalam peramalan suhu udara menggunakan 80% data *training* dan 20% data *testing*. Adapun model LSTM untuk peramalan Suhu Udara Minimum dibentuk dari 50 *neuron hidden*, 4 *batch size*, dan 50 *epoch* dengan nilai RMSE 0,89 dan MAPE sebesar 2,59%. Sedangkan model LSTM terbaik untuk peramalan Suhu Udara Maksimum dan Suhu Udara Rata-rata dibentuk dari 25 *neuron hidden*, 4 *batch size*, dan 50 *epoch* dengan nilai RMSE dan MAPE Suhu Udara maksimum sebesar 0,85 dan 2,32% serta Suhu Udara Rata-rata sebesar 0,88 dan 2,44%. Maka model LSTM ini sangat baik untuk peramalan Suhu Udara.

Kata Kunci: Peramalan, *Time Series*, *Long Short Term Memory* (LSTM).

**PEMODELAN *TIME SERIES* UNTUK PERAMALAN SUHU UDARA
MENGUNAKAN METODE *LONG SHORT TERM MEMORY (LSTM)*
(Studi Kasus: Stasiun Klimatologi Lampung)**

Oleh

LINDA SUNDARI

Skripsi

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

Pada

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



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

Judul Skripsi :

**PEMODELAN *TIME SERIES* UNTUK
PERAMALAN SUHU UDARA
MENGUNAKAN METODE *LONG SHORT
TERM MEMORY (LSTM)*
(Studi Kasus : Stasiun Klimatologi Lampung)**

Nama Mahasiswa :

Linda Sundari

Nomor Pokok Mahasiswa :

1917031070

Jurusan :

Matematika

Fakultas :

Matematika dan Ilmu Pengetahuan Alam



1. **Komisi Pembimbing**

Drs. Nusyirwan, M.Si.
NIP 196610101992031028

Pandri Ferdias, S.Si., M.Sc.
NIP 198706122019031007

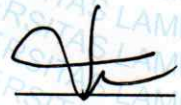
2. **Ketua Jurusan Matematika**

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Dr. Nusyirwan, M.Si.**



Sekretaris : **Pandri Ferdias, S.Si., M.Sc.**



Penguji
Bukan Pembimbing : **Prof. Drs. Mustofa Usman, M.A., Ph.D.**



2. Plt. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung.

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



Tanggal Lulus Ujian Skripsi: **20 Februari 2023**



PERNYATAAN SKRIPSI MAHASISWA

Nama : **Linda Sundari**
Nomor Pokok Mahasiswa : **1917031070**
Jurusan : **Matematika**
Judul Skripsi : **Pemodelan *Time Series* untuk Peramalan Suhu Udara Menggunakan Metode *Long Short Term Memory* (LSTM)
(Studi Kasus: Stasiun Klimatologi Lampung)**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila dikemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 20 Februari 2023
Yang Menyatakan



Linda Sundari
NPM. 1917031070

RIWAYAT HIDUP

Penulis dilahirkan di Desa Labung Nala, Kabupaten Lampung Selatan pada Tanggal 18 Juni 1999, sebagai anak kedua dari empat bersaudara, dari pasangan Bapak Mastur dan Ibu Neneng Maya Kariyati.

Penulis memulai pendidikan Sekolah Dasar di SD N 1 Lebung Nala. Penulis melanjutkan Sekolah Menengah Pertama di MTs N Singkut pada Tahun 2012-2015, dan Sekolah Menengah Atas di SMA N 2 Sarolangun Kabupaten Sarolangun Provinsi Jambi.

Penulis melanjutkan jenjang pendidikan di Universitas Lampung melalui jalur SBMPTN serta sebagai penerima Beasiswa Bidikmisi tahun 2019. penulis terdaftar sebagai Mahasiswa Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Pada tahun 2022, penulis melakukan Kuliah Praktik (KP) di Dinas Ketahanan Pangan, Tanaman Pangan dan Hortikultura kemudian dilanjutkan Kuliah Kerja Nyata (KKN) di Desa Nibung, Kecamatan Gunung Pelindung, Kabupaten Lampung Timur. Selama perkuliahan, penulis mengikuti beberapa organisasi yaitu Badan Eksekutif Mahasiswa (BEM) FMIPA sebagai Staff Ahli Dinas Sains (2020-2021) dan Pengabdian Masyarakat, Rohani Islam (ROIS) FMIPA sebagai Anggota Bidang Akademik dan Riset (2020).

KATA INSPIRASI

“Kemudian apabila kamu telah membulatkan tekad, maka bertawakallah kepada Allah Subhaanahu Wata;aalaa. Sesungguhnya Allah Subhaanahu Wata;aalaa menyukai orang yang bertawakkal kepada-Nya”

Q.S. Ali Imran : 159

Seorang pemuda tidak akan sia-sia kecuali dengan empat perkara: agama, amanah, menjaga diri, dan kesungguhan

Imam Syafi'i

“Kamu terus mengeluh tapi tetap dikerjakan, itu adalah sikap profesionalisme dalam hidup”

Tri Guna

Jika kamu tidak berjuang untuk apa yang kamu inginkan, jangan menangis untuk apa yang tidak kamu dapatkan”

PERSEMBAHAN

Puji dan syukur penulis ucapkan kepada Allah Subhaanahu Wata;aalaa atas segala rahmat dan karunia-Nya. Tak lupa salawat beserta salam senantiasa tercurahkan kepada Nabi Muhammad Shalallaahu Alaihi Wassalaam yang merupakan suri tauladan terbaik bagi seluruh umat.

Kupersembahkan skripsi ini sebagai karya yang telah penulis buat kepada:

Kedua Orang Tua Tercinta

Terima kasih atas segala hal yang telah kalian berikan, baik dari segi dukungan materi, mental dan do'a yang tidak pernah terputus. Semoga kebaikan Bapak dan Ibu menjadikan amalan yang terbaik di sisi Allah Subhaanahu Wata;aalaa dan selalu dimuliakan oleh-Nya serta selalu diberikan keberkahan hidup.

Kakak tersayang

Karya ini tidak lepas dari dukungan, saran, doa dan motivasi yang selalu kakak berikan selama ini. Semoga Allah Subhaanahu Wata;aalaa menjadikan semua kebaikan ini untuk selalu diberikan keberkahan selama hidupnya.

Teman Terbaik

Karya ini merupakan wujud dari dukungan dan doa yang selalu kalian berikan. Tidak ada kata lain yang layak selain ucapan Terima kasih. Semoga hubungan ini terus berlanjut dan terus terjaga dalam kebaikan.

Almamater Kebanggaan Universitas Lampung

SANWACANA

Puji syukur penulis ucapkan kehadiran Allah Subhaanahu Wata;aalaa, karena berkat rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan tepat waktu. Skripsi dengan judul “Pemodelan *Time Series* untuk Peramalan Suhu Udara Menggunakan Metode *Long Short Term Memory* (LSTM) (Studi Kaus: Stasiun Klimatologi Lampung)” merupakan salah satu syarat untuk memperoleh gelar sarjana dMatematika di Universitas Lampung.

Pada kesempatan ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Drs. Nusyirwan, M.Si., selaku pembimbing I yang telah memberikan masukan, saran serta bimbingan kepada penulis selama proses penyelesaian skripsi ini.
2. Bapak Pandri Ferdias, S.Si., M.Sc., selaku pembimbingan II yang telah memberikan masukan, saran serta bimbingan kepada penulis selama proses penyelesaian skripsi.
3. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku dosen pembahas yang telah memberikan evaluasi, saran dan kritikan hingga skripsi ini dapat terselesaikan.
4. Ibu Dr. Notiragayu, S.Si., M.Si., selaku pembimbing akademik yang telah membantu dalam memberikan bimbingan kepada penulis selama perkuliahan ini dilakukan.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Universitas Lampung.

6. Bapak Dr. Eng. Heri Satria, S.Si. M.Si., selaku Plt. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh civitas akademik, dosen, serta staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Ibu, Bapak, Saudaraku serta seluruh keluarga besar yang selalu memberikan dukungan dan doa yang terbaik agar penulis diberikan kelancaran serta kemudahan dalam menyelesaikan skripsi ini.
9. Teman-teman terdekat; Berlian, Coi, Mega, Tika dan Zida yang selalu memberikan bantuan, dukungan, masukan, serta kebersamaan dan pembelajaran hidup selama perkuliahan ini berlangsung.
10. Teman-teman saya; iqi dan sinta sudah memberikan bantuan, dukungan dan masukan dalam menyelesaikan skripsi saya.
11. Seluruh teman-teman mahasiswa Jurusan Matematika angkatan 2019.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat dibutuhkan, guna penyempurnaan skripsi ini.

Bandar Lampung, 20 Februari 2023
Penulis,

Linda Sundari

DAFTAR ISI

	Halaman
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvi
I. PENDAHULUAN	1
1.1. Latar Belakang dan Masalah	1
1.2. Tujuan Penelitian	4
1.3. Manfaat Penelitian	4
II. TINJAUAN PUSTAKA	6
2.1. Peramalan (<i>Forecasting</i>).....	6
2.2. <i>Time Series</i>	7
2.3. <i>Preprocessing</i>	8
2.3.1. <i>Data Cleaning</i>	8
2.3.2. Normalisasi Data.....	9
2.3.3. Proses Segmentasi Data.....	9
2.3.4. Pembagian Data <i>Training</i> dan Data <i>Testing</i>	10
2.4. <i>Artificial Neural Network</i> (ANN).....	10
2.4.1. Arsitektur <i>Artificial Neural Network</i> (ANN).....	11
2.4.2. Proses Pembelajaran <i>Artificial Neural Network</i> (ANN)	13
2.5. <i>Deep Learning Neural Network</i>	14
2.6. <i>Recurrent Neural Network</i> (RNN)	15
2.7. <i>Long Short Term Memory</i> (LSTM)	16
2.7.1 Arsitektur <i>Long Short Term Memory</i> (LSTM)	16
2.7.2 Bias dan <i>Weight</i>	21
2.8. Denormalisasi.....	21

2.9.	Evaluasi Model.....	22
2.9.1	<i>Root Mean Square Error (RMSE)</i>	22
2.9.2	<i>Mean Absolute Percentage Error (MAPE)</i>	23
III.	METODOLOGI PENELITIAN.....	24
3.1	Waktu dan Tempat Penelitian	24
3.2	Data Penelitian	24
3.3	Metode Penelitian	25
IV.	HASIL DAN PEMBAHASAN.....	27
4.1.	Dataset Suhu Udara	27
4.2.	<i>Preprocessing</i>	28
4.2.1.	<i>Data Cleaning</i>	28
4.2.2.	Normalisasi Data	33
4.2.3.	Segmentasi Data	35
4.2.4.	Pembagian Data <i>Training</i> dan Data <i>Testing</i>	36
4.3.	Building <i>Model Long Short Term Memory (LSTM)</i>	37
4.4.	Perhitungan Manual.....	37
4.5.	Pengujian Model.....	45
4.6.	Validasi Model	48
4.7.	Peramalan Suhu Udara	53
V.	KESIMPULAN.....	58
	DAFTAR PUSTAKA.....	59
	LAMPIRAN.....	62

DAFTAR TABEL

Tabel	Halaman
1. Kategori <i>Range</i> Nilai MAPE.....	23
2. Dataset suhu udara	28
3. Jumlah <i>missing value</i> pada data suhu udara.....	30
4. Data suhu udara setelah dilakukan data <i>cleaning</i>	31
5. Nilai minimum dan maksimum data suhu udara	33
6. Hasil normalisasi suhu udara.....	34
7. Pembagian data <i>training</i> dan data <i>testing</i>	36
8. Nilai <i>input</i> , <i>weight</i> , dan bias pada <i>timestep</i> pertama.....	37
9. Nilai prediksi suhu udara.....	45
10. Data prediksi suhu udara setelah denormalisasi	47
11. Nilai akurasi model LSTM berdasarkan jumlah <i>neuron hidden</i>	49
12. Nilai akurasi model LSTM berdasarkan jumlah <i>batch size</i>	49
13. Nilai akurasi model LSTM berdasarkan jumlah <i>epoch</i>	50
14. Nilai akurasi model LSTM berdasarkan pembagian <i>training</i> dan <i>testing</i>	51
15. Hasil Peramalan Suhu Udara.....	54

DAFTAR GAMBAR

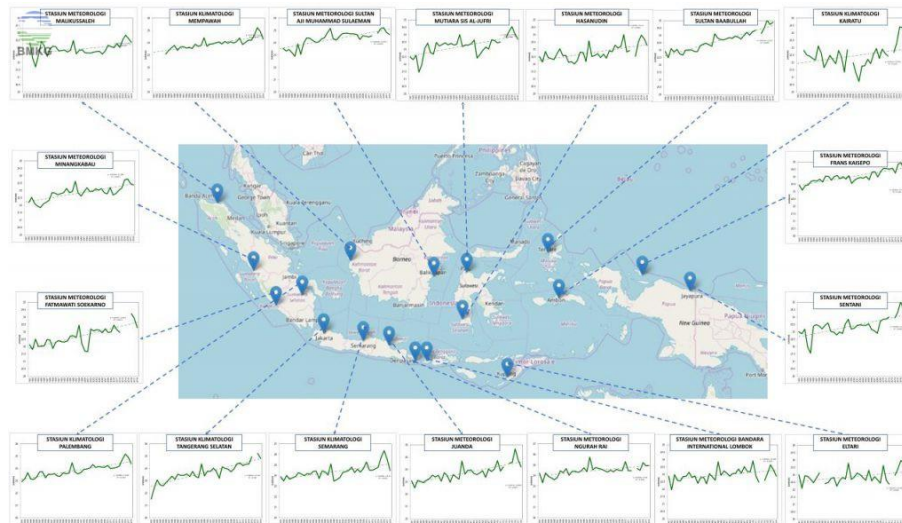
Gambar	Halaman
1. Perubahan suhu udara di berbagai Stasiun BMKG	2
2. Ilustrasi pola <i>time series</i>	9
3. Arsitektur <i>Artificial Neural Network</i> (ANN)	11
4. Arsitektur <i>Single Layer Network</i>	12
5. Arsitektur <i>Multi Layer Network</i>	13
6. Arsitektur <i>Recurrent Network</i>	13
7. Arsitektur <i>Deep Learning Neural Network</i> (DLNN).....	15
8. Arsitektur <i>Long Short Term Memory</i> (LSTM).....	17
9. Diagram alir metode penelitian	26
10. Grafik <i>time series</i> suhu udara	27
11. Plot data <i>cleaning</i> suhu udara.....	32
12. Ilustrasi pola <i>time series</i>	35
13. Plot prediksi suhu udara berdasarkan data <i>testing</i>	48
14. Model LSTM Suhu Udara Minimum.....	52
15. Model LSTM Suhu Udara Maksimum dan Suhu Udara Rata-Rata	53
16. Plot peramalan suhu udara	53

I. PENDAHULUAN

1.1. Latar Belakang

Forecasting (peramalan) adalah pengetahuan mengenai perencanaan dan pengendalian terkait dengan kejadian pada masa yang akan datang. peramalan dilakukan dengan menyertakan data historis yaitu data *time series* dan memproyeksikan dengan model-model matematis ke masa yang akan datang. Data *time series* atau data deret waktu merupakan data yang dikumpulkan pada periode waktu tertentu. Misalnya, data yang diambil berdasarkan periode harian, periode mingguan, periode bulanan ataupun data yang diambil dalam periode tahunan. Peramalan ini dapat digunakan dalam pengambilan keputusan untuk mencapai tujuan tertentu. Mengingat betapa pentingnya peramalan ini, banyak berbagai studi mengenai peramalan telah dilakukan, diantaranya peramalan mengenai harga suatu barang, harga saham, kunjungan wisatawan maupun cuaca.

Cuaca memiliki peran yang penting untuk kehidupan kita sehari-hari. Cuaca memiliki beberapa unsur yang penting, salah satunya yaitu suhu udara. Unsur cuaca ini mengalami perubahan yang terus meningkat dalam beberapa tahun terakhir. Hal ini dapat diperhatikan pada Gambar 1 berikut.



Gambar 1. Perubahan suhu udara di beberapa Stasiun BMKG
(Sumber: Badan Meteorologi, Klimatologi dan Geofisika)

Berdasarkan Gambar 1, terlihat adanya tren suhu udara yang terus mengalami peningkatan pada beberapa Stasiun BMKG. Perubahan ini merupakan permasalahan yang dapat mempengaruhi berbagai hal, seperti permasalahan mengenai lingkungan, kesehatan, pertanian, hingga menyebabkan permasalahan dalam perencanaan pembangunan. Oleh karena itu, peramalan suhu udara merupakan perkara penting yang perlu diperhatikan..

Dalam perkembangan era digital ini, teknologi informasi juga terus mengalami peningkatan yang pesat, seperti berkembangnya pengetahuan kecerdasan buatan atau *artificial intelligence* (AI) yaitu jaringan syaraf tiruan. Peramalan suhu udara merupakan penerapan dari sains dan teknologi dalam memprediksi keadaan atmosfer pada waktu dan lokasi tertentu. Biasanya, suhu udara ditampilkan oleh BMKG dalam bentuk data deret waktu (*time series*) berupa data harian. Studi kasus mengenai peramalan suhu udara telah banyak dilakukan dengan berbagai metode peramalan, diantaranya peramalan suhu udara menggunakan metode analisis fourier (Septiyani, dkk., 2019), metode *Autoregressive Integrated Moving Average* (ARIMA) dan Kalman Filter (Ulinuha & Farida, 2018), *Singular Spectrum Analysis* dan *Fourier Series Analysis* (Christienova, 2018),

metode *fuzzy time series* dan Markov Chain (Wulandari, 2021), Algoritma *Support Vector Machine* (Rusdi, 2017), Metode LSTM (Khumaidi, dkk., 2020) serta metode peramalan lainnya.

Selain digunakan untuk analisis peramalan suhu udara, terdapat beberapa penelitian terdahulu yang menerapkan LSTM. Penelitian tersebut diantaranya yaitu prediksi curah hujan Kota Malang (Rizki, dkk., 2020), prediksi penjualan produk PT. Mestika Farma (Wiranda & Sadikin, 2019), analisis peramalan kasus Covid-19 di Jawa Tengah (Qori, dkk., 2022), Prediksi pendapatan kargo (Aprian, dkk., 2020), prediksi Harga Bitcoin (Aldi, dkk., 2018), dan sebagainya.

Penelitian mengenai metode LSTM ini mengacu pada penelitian Khumaida, dkk. (2020) mengenai penerapan algoritma LSTM dalam memprediksi unsur cuaca yang meliputi kualitas udara serta suhu udara di Kota Bandung. Pada penelitian tersebut, model LSTM dibentuk berdasarkan data *time series* yang menggunakan *hidden layer* sebanyak 4 layer, 32 *batch size*, 1000 *epoch* dan menggunakan optimizer adam. Penelitian tersebut menghasilkan model LSTM yang cukup baik dalam memprediksi kualitas udara dan suhu udara. Adapun nilai RMSE yang dihasilkan untuk metode LSTM dalam memprediksi kualitas suhu udara yaitu 1,85 serta nilai RMSE pada suhu udara sebesar 3,15 (Khumaida, dkk., 2020).

Selain itu, penelitian mengenai metode LSTM ini juga mengacu pada penelitian Rizki, dkk. (2020) mengenai penerapan *Deep Learning* pada arsitektur LSTM dalam memprediksi curah hujan di Kota Malang. Pada penelitian tersebut, peneliti melakukan perbandingan model LSTM dengan memberikan berbagai nilai parameter LSTM (*neuron hidden*, *batch*, *epoch*). Model LSTM dalam memprediksi curah hujan Kota Malang yang memiliki nilai RMSE yang terkecil yaitu model LSTM dengan *neuron hidden* sebesar 256 dan *epoch* sebesar 150 (Rizki, dkk., 2020).

Adapun penelitian mengenai peramalan suhu udara dengan studi kasus Stasiun Klimatologi Lampung baru pertama kali dilakukan. Oleh sebab itu, penulis melakukan penelitian mengenai pemodelan *time series* dalam prediksi suhu udara menggunakan metode *Long Short Term Memory* (LSTM) pada studi kasus Stasiun Klimatologi Lampung. Variabel suhu udara yang akan dilakukan analisis meliputi suhu udara minimum (X1), suhu udara maksimum (X2) dan suhu udara rata-rata (X3) dengan membandingkan model LSTM yang dibentuk dari berbagai nilai parameter dalam model LSTM yaitu *neuron hidden*, *batch size*, dan *epoch*. Model LSTM yang mempunyai nilai MAPE dan RMSE terkecil merupakan model LSTM yang terbaik dalam memprediksi suhu udara tersebut.

1.2. Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini yaitu:

1. Mengetahui model LSTM berdasarkan data *time series* pada peramalan suhu udara di Stasiun Klimatologi Lampung.
2. Mengetahui performa metode *Long Short Term Memory* dalam meramalkan suhu udara.
3. Mengetahui arsitektur jaringan terbaik yang didapatkan dari metode *Long Short Term Memory* pada peramalan suhu udara.
4. Mengetahui hasil peramalan suhu udara berdasarkan data *time series* suhu udara di Stasiun Klimatologi Lampung yang dihasilkan dari model LSTM yang terbaik.

1.3. Manfaat Penelitian

Penelitian ini merupakan bahan analisis pemodelan *time series* untuk peramalan data suhu udara dengan studi kasus Stasiun Klimatologi Lampung menggunakan

data historis dari tanggal 1 Januari 2018 hingga 30 Oktober 2022 yang berupa data harian. Selain itu, penelitian ini juga memiliki manfaat yaitu:

1. Meningkatkan wawasan mengenai implementasi metode LSTM dalam peramalan suhu udara.
2. Menjadi sarana pengembangan keilmuan bagi mahasiswa dan masyarakat umum khususnya di bidang peramalan suhu udara.

II. TINJAUAN PUSTAKA

2.1. Peramalan (*Forecasting*)

Peramalan atau disebut juga *forecasting* adalah ilmu yang mempelajari tentang memperkirakan suatu hal berdasarkan data-data pada masa lampau (Qori, dkk., 2022). Peramalan dianalisis berdasarkan data historis yaitu data deret waktu (*time series*) sehingga dapat digunakan dalam memprediksi suatu hal yang akan datang. Hal ini merupakan perihal penting dalam banyak hal yaitu meliputi permasalahan mengenai lingkungan, kesehatan, pertanian, hingga permasalahan dalam perencanaan pembangunan.

Pendekatan dalam metode peramalan meliputi dua jenis pendekatan, yaitu pendekatan kualitatif dan pendekatan kuantitatif. Pendekatan kualitatif merupakan pendekatan peramalan berdasarkan opini atau pendapat pihak terkait dikarenakan tidak adanya suatu informasi data historis sebelumnya sebagai dasar perkiraan. Misalnya pengenalan pengenalan suatu produk yang baru pertama kali dikeluarkan sehingga belum ada riwayat yang berhubungan dengan hal tersebut. Sedangkan metode peramalan dengan pendekatan kuantitatif adalah pendekatan dalam peramalan yang mengacu pada data historis sebagai dasar perkiraan (Montgomery, 2008).

Menurut Montgomery (2008), Proses peramalan meliputi:

1. Definisi Masalah (*Problem definition*): pengembangan pemahaman tentang bagaimana ramalan akan dilakukan dan digunakan.

2. Pengumpulan data (*Data collection*): mengumpulkan data yang berhubungan dengan variabel yang akan diramal.
3. Menganalisis data (*Data analysis*): mengenali pola data yang akan digunakan dalam peramalan. Hal tersebut dapat dilakukan dengan membangun plot data deret waktu dan dilakukan pemeriksaan secara visual pada plot data yang diperoleh.
4. Pemilihan serta pemasangan model (*Model selection and fitting*): menentukan metode peramalan yang sesuai dengan pola data yang diambil.
5. Evaluasi Model (*Model evaluation*): mengevaluasi performa dari model peramalan yang dibentuk yaitu apakah suatu model tersebut memiliki kinerja yang baik atau tidak dalam melakukan analisis peramalan pada data yang digunakan.
6. Menampilkan hasil peramalan: model yang telah dipilih digunakan dalam melakukan peramalan untuk periode yang akan datang.
7. *Monitoring forecasting model performance*: memantau kinerja model peramalan untuk memastikan performa kinerjanya dalam aplikasi yang dimaksud.

2.2. Time Series

Time series adalah data yang diambil berdasarkan periode tertentu guna memperkirakan suatu kegiatan pada waktu tertentu. Analisis *time series* dapat menentukan bagaimana suatu kejadian berhubungan dengan kejadian lainnya (Qori, dkk., 2022). Pengukuran data *time series* dilakukan secara berurutan (Afriansyah, 2017).

2.3. Preprocessing

Preprocessing data merupakan proses persiapan data yang telah diperoleh sebelum dilakukan proses pelatihan . Proses ini perlu dilakukan sehingga data yang diperoleh dapat digunakan dalam pelatihan dan bekerja secara optimal (Aprian, dkk., 2020).

2.3.1. Data Cleaning

Data cleaning merupakan teknik yang digunakan untuk penanganan *missing value* atau *noise*. *Missing value* merupakan kondisi dimana suatu dataset yang diambil memiliki data yang hilang atau tidak lengkap. *Data cleaning* untuk penanganan *missing value* dapat dilakukan dengan dua cara yaitu (Aprian, dkk., 2020):

1. Menghapus keseluruhan baris (*row*) pada data yang hilang
2. Menghitung nilai pengganti (*imputation*) dengan nilai konstanta, nilai rata-rata, nilai median, atau mode untuk kolom yang mengalami missing atau beberapa imputasi lain

Adapun persamaan nilai rata-rata (*mean*) dari suatu data dirumuskan dalam persamaan berikut:

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

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

2.3.2. Normalisasi Data

Normalisasi data atau *scaling* data merupakan proses transformasi data yang mengubah data asli ke dalam bentuk lain. Proses normalisasi perlu dilakukan agar

menghasilkan nilai *error* yang sekecil mungkin. Teknik normalisasi yang digunakan untuk peramalan suhu udara ini adalah *MinMaxScaler* atau bisa disebut juga dengan *Min-Max Scaling*. Teknik ini mentransformasikan data aktual ke dalam nilai dengan rentang [0,1]. Teknik *Min-Max Scaling* dirumuskan ke dalam persamaan berikut (Rizki, dkk., 2020):

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

Keterangan:

x' : Data suhu udara yang telah dinormalisasikan.

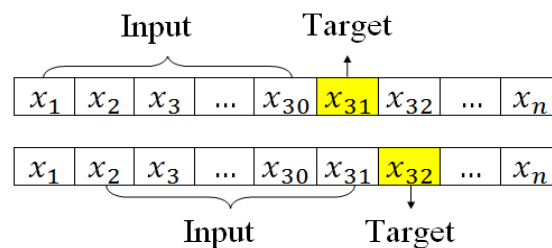
x : Data aktual suhu udara.

X_{min} : Nilai minimum data suhu udara.

X_{max} : Nilai maksimum data suhu udara.

2.3.3. Proses Segmentasi Data

Segmentasi adalah proses memisahkan serta mengelompokkan data menjadi data yang diperlukan oleh sistem. Segmentasi data dilakukan dengan pembentukan pola *time series* pada data. Proses segmentasi untuk penelitian ini diilustrasikan pada Gambar 2 berikut.



Gambar 2. Ilustrasi pola *time series*

Pada Gambar 2, apabila dilakukan pengelompokan yang terdiri dari 30 data, maka data ke-1 hingga data ke-30 merupakan bagian dari *input*. Sedangkan

bagian target merupakan data ke-31. Begitu pula seterusnya, apabila *input* terdiri dari data ke-2 hingga data ke-31, maka targetnya merupakan data ke-32.

2.3.4. Pembagian Data *Training* dan Data *Testing*

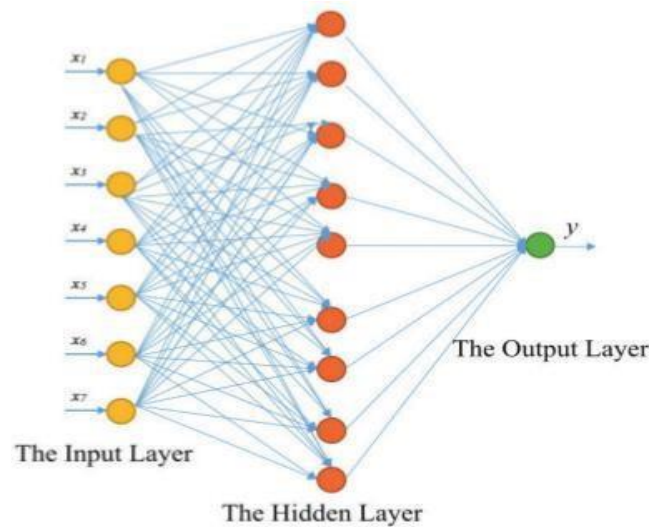
Pembagian data merupakan tahap membagi suatu dataset ke dalam dua bagian, yaitu data untuk *training* dan data untuk *testing*. Data untuk *training* adalah bagian dari dataset yang dipakai untuk pelatihan suatu model. Sedangkan, data untuk *testing* adalah dataset yang dipakai sebagai pengujian keakuratan atau performa dari model yang dibentuk dari pelatihan sebelumnya. Pembagian data ini bertujuan untuk mendapatkan model yang memiliki akurasi sebaik mungkin (Cahyadi, dkk., 2020).

2.4. Artificial Neural Network (ANN)

Artificial Neural Network (ANN) merupakan sistem komputerisasi sebagai pemrosesan data yang memiliki sistem kerja yang sama dengan jaringan syaraf atau sistem kerja otak pada manusia. ANN memiliki sistem kerja yang sama dengan otak pada manusia dimana mempunyai neuron-neuron yang saling berkaitan satu sama lain dalam berbagai lapisan jaringan.

2.4.1. Arsitektur *Artificial Neural Network* (ANN)

Arsitektur ANN memiliki elemen penting yang disebut *layer* (lapisan). Pada arsitektur ini, *layer* digunakan sebagai pemrosesan matematis guna memahami informasi yang diberikan. Arsitektur dari ANN diilustrasikan pada Gambar 3. berikut.



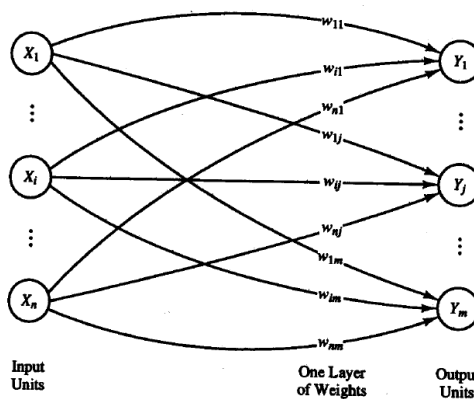
Gambar 3. Arsitektur *Artificial Neural Network* (ANN)
(Sumber: Matondang, 2013)

Adapun layer pada arsitektur ANN dibagi menjadi tiga *layer*, yaitu:

1. *Input layer* : berperan sebagai input terima dimana data yang ingin diproses dimasukkan ke dalam arsitektur ANN. Pada *layer* ini, tersimpan informasi-informasi mengenai data yang diberikan dan data-data akan disalurkan menuju lapisan *layer* selanjutnya.
2. *Hidden layer* : berperan sebagai penghubung antara lapisan *input* dan *output layer*. Pada arsitektur ANN, data yang berasal dari *layer* sebelumnya yaitu *input layer* dapat melewati lebih dari satu *hidden layer*. Pada *layer* ini dilakukan proses perhitungan untuk menemukan fitur dan pola yang tersembunyi.
3. *Output layer* : *layer* yang berperan sebagai penerima hasil perhitungan dan menampilkan hasil perhitungan dari *layer* sebelumnya menggunakan fungsi aktivasi.

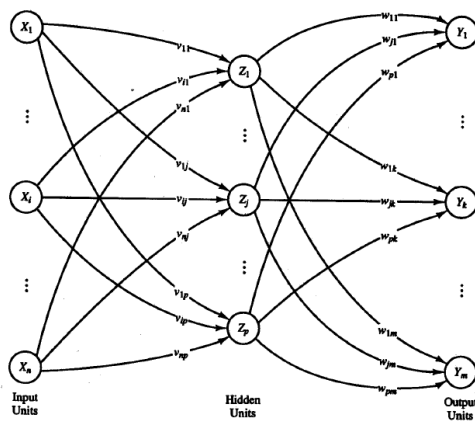
Arsitektur ANN terdiri dari *layer-layer* serta neuron untuk setiap layernya yang dihubungkan oleh bobot (*weight*), fungsi aktivasi dan *learning function*. ANN mempunyai 3 pemodelan jaringan, yaitu: (Matondang, 2013)

1. *Single layer network*, merupakan pemodelan pada arsitektur ANN yang mempunyai satu *input layer* dan satu *output layer*. Cara kerja jaringan ini adalah sumber *input layer* yang berupa node diproyeksikan menuju *output layer* yang berasal neuron. Pada jaringan ini, pengolahan data hanya melalui dua *layer*, yaitu penerimaan inputan dan pengolahan data dilakukan oleh *input layer* lalu mengirimkan hasilnya pada *output layer*.



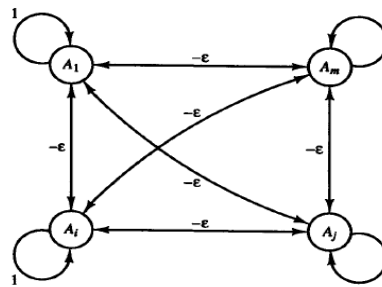
Gambar 4. Arsitektur *Single Layer Network*
(Sumber: Fausset, 1994)

2. *Multi layer Network*, merupakan arsitektur ANN yang memiliki tiga jenis layer yaitu *input layer*, *hidden layer* dan *output layer*. Pemodelan ini lebih baik digunakan dalam penyelesaian permasalahan yang lebih kompleks dibandingkan pemodelan *single layer*. Sistem kerja pemodelan ini dimulai dengan *input layer* yang mengirimkan informasi data pada jaringan, kemudian melakukan komputasi menuju *layer* selanjutnya. Adapun *output* yang diperoleh sebelumnya akan digunakan sebagai inputan dari *layer* selanjutnya.



Gambar 5. Arsitektur *Multi Layer Network*
(Sumber: Fausset, 1994)

3. *Recurrent network*, yaitu jaringan ANN yang digunakan dalam memproses suatu data yang memiliki urutan (*sequential data*).



Gambar 6. Arsitektur *Recurrent Network*
(Sumber: Fausset, 1994)

2.4.2. Proses Pembelajaran *Artificial Neural Network* (ANN)

ANN memiliki proses pembelajaran yang terdiri dari dua bentuk, yaitu *Supervised Learning* dan *Unsupervised Learning* (Sari, 2016).

1. *Supervised Learning*

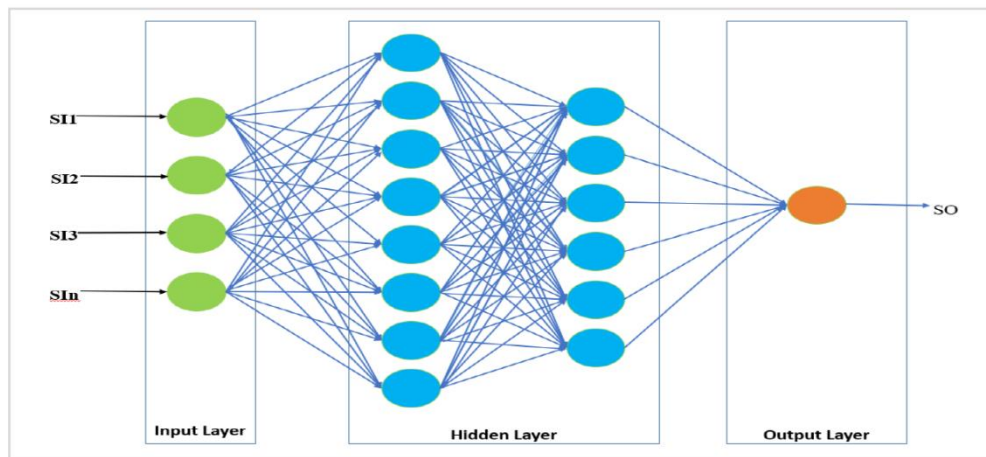
Supervised learning atau disebut juga pembelajaran yang terawasi merupakan ANN dengan proses pembelajaran yang melibatkan data sebelumnya. Proses *learning* ini mempunyai *output* dengan target yang sudah diketahui sebelumnya. Adapun cara melakukan pelatihan dalam pembelajaran ini yaitu memberikan data untuk pelatihan yang mencakup *input* dan *output* yang diinginkan.

2. *Unsupervised Learning*

Unsupervised Learning atau disebut juga pembelajaran tidak terawasi merupakan ANN dengan proses pembelajaran yang tidak membutuhkan target sebelumnya. Pembelajaran ini hanya membutuhkan data *input* saja dan tidak memerlukan adanya target. Beberapa metode yang menggunakan proses pembelajaran ini yaitu *Kohonen Self-organizing Maps* dan *Counterpropagation*.

2.5. *Deep Learning Neural Network*

Menurut Supriyadi (2020), *Deep Learning Neural Network* merupakan salah satu pengembangan dari *machine learning* yang memanfaatkan algoritma *Artificial Neural Network* (Jaringan syaraf tiruan). *Deep Learning Neural Network* memiliki arsitektur yang kuat untuk *supervised learning* dan memiliki jumlah *hidden layer* lebih dari satu. Dengan lebih banyak lapisan ini, maka model pembelajaran *Deep Learning Neural Network* bisa di implementasikan pada dataset yang cukup besar dengan kinerja yang lebih baik. Akan tetapi, metode ini mempunyai kekurangan dalam hal kecepatan analisis. Apabila semakin banyak menggunakan layer pada proses pengolahan, maka semakin lama proses pengolahannya. *Deep learning* dapat digunakan dalam hal peramalan data *time series*, salah satunya dengan menggunakan metode LSTM (Pratiwi, dkk., 2021). Berikut ini ilustrasi dari arsitektur DLNN.



Gambar 7. Arsitektur *Deep Learning Neural Network* (DLNN)
(Sumber: Rizki, dkk., 2020)

2.6. *Recurrent Neural Network* (RNN)

Recurrent Neural Network (RNN) adalah metode yang dirancang untuk menganalisis data *sequential* dengan memanfaatkan algoritma *deep learning*. RNN merupakan kelas dalam *deep learning* dengan proses pembelajaran yang terawasi. Dikarenakan kemampuan RNN dalam analisis data *sequential*, maka metode ini dapat digunakan untuk memproses data *time series*. Kelas *deep learning* ini memiliki arsitektur yang sama dengan *deep learning* yaitu *input layer*, *hidden layer*, dan *output layer* (Larasati, 2020). Berdasarkan arsitektur yang dimiliki RNN, secara teori metode ini dapat menyelesaikan permasalahan ketergantungan dalam jangka panjang. Akan tetapi, dalam penggunaannya RNN tidak mampu melakukan penyimpanan informasi sebelumnya dengan baik dikarenakan permasalahan gradien yang menghilang.

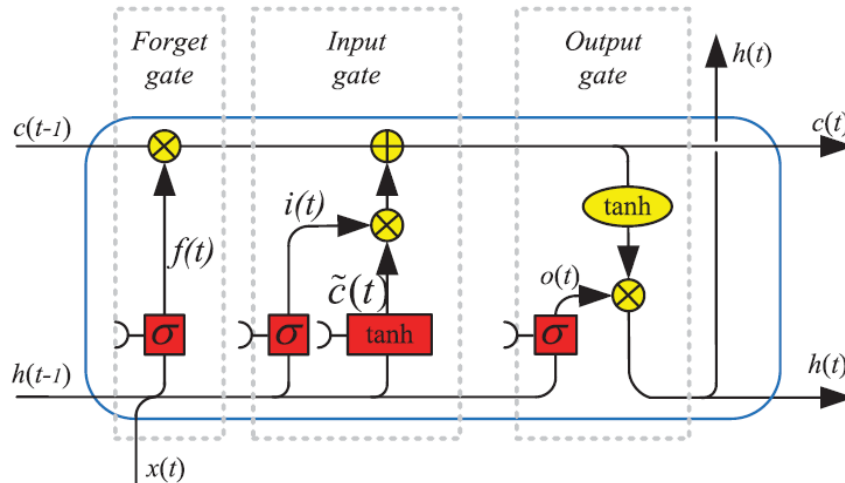
2.7. Long Short Term Memory (LSTM)

Pada tahun 1997, pengembangan algoritma *deep learning* telah dilakukan dan merupakan suatu modifikasi dari metode RNN yang dikenal dengan Long Short Term Memory (LSTM). Metode ini dikembangkan oleh Hochreiter dan Schmidhuber. Pengembangan metode ini terus dilakukan dalam berbagai bidang yaitu dalam bidang peramalan (*forecasting*) dan bidang *speed recognition*. LSTM mampu mengatasi kekurangan dari RNN yang tidak dapat melakukan prediksi berdasarkan informasi pada masa lalu yang disimpan untuk jangka waktu yang lama. Hal ini dikarenakan seiring berjalannya waktu informasi sebelumnya akan terganti dengan memori baru yang berakibat berkurangnya akurasi dari prediksi yang dilakukan oleh RNN (Zhao, *et al.*, 2017).

LSTM mengatasi kekurangan dari RNN karena metode ini memiliki *memory cell* dan *gate units* yang mampu menyimpan informasi setiap masukan dalam jangka waktu yang panjang, meskipun data yang digunakan sangat besar (Karpathy, *et al.*, 2015).

2.7.1. Arsitektur Long Short Term Memory (LSTM)

Arsitektur LSTM memiliki tiga gerbang utama, diantaranya adalah *input gate*, *forget gate*, dan *ouput gate*. *Memory cell* beserta setiap gerbang dapat melakukan pembacaan, penyimpanan dan perbaruan informasi sebelumnya (Hochreiter & Schmidhuber, 1997).



Gambar 8. Arsitektur *Long Short Term Memory* (LSTM)
(Sumber: Yu, *et al.*, 2019)

1. *Forget Gate* (f_t)

Forget gate adalah proses menghilangkan informasi yang tidak dibutuhkan dengan menggunakan fungsi sigmoid. Pada *gate* ini, fungsi sigmoid mentransformasikan nilai antara -1 dan 1 menjadi nilai antara 0 dan 1. Fungsi sigmoid dirumuskan dalam persamaan berikut (Buduma, 2017).

$$\text{Sigmoid}(x) = \sigma(x) = \frac{1}{1+e^{-x}} \quad (2.4)$$

Adapun turunan dari fungsi sigmoid dinyatakan dalam persamaan berikut.

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \quad (2.5)$$

Selanjutnya, perhitungan *forget gate* dilakukan dengan rumus:

$$f_t = \sigma(W_{fx} \cdot x_t + W_{fh} \cdot h_{t-1} + b_f) \quad (2.6)$$

dimana:

- f_t : *Forget gate* pada *timestep* ke-t
- W_f : Nilai *weight* (bobot) untuk *forget gate*
- x_t : Nilai data *input* suhu udara untuk *timestep* ke-t.
- h_{t-1} : Nilai *output* pada *timestep* sebelumnya (t-1)
- b_f : Nilai bias pada *forget gate*

Pada bagian *gate* ini dilakukan perhitungan berapa banyak c_{t-1} pada *timestep* sebelumnya dapat dicadangkan ke dalam *cell state* pada *timestep* saat ini. Adapun gerbang ini memunculkan perhitungan dengan nilai diantara 0 dan 1.

2. *Input gate* (i_t)

Input gate adalah proses mengolah informasi diolah untuk menentukan informasi yang diperbarui dan kemudian dicadangkan ke dalam status sel C_t . *Gate* ini mempunyai fungsi utama yaitu mendeteksi *cell* yang perlu diperbarui memperbarui informasi ke tahapan selanjutnya. Dalam *input gate*, nilai yang akan diperbarui dipilih oleh fungsi sigmoid yang didefinisikan sebagai berikut.

$$i_t = \sigma(W_{ix} \cdot x_t + W_{ih} \cdot h_{t-1} + b_i) \quad (2.7)$$

dimana:

- i_t : Nilai *Input gate* untuk *timestep* ke-t
- σ : Fungsi sigmoid
- W_i : Bobot pada *input gate*
- x_t : Nilai *input data* suhu udara untuk *timestep* ke-t
- h_{t-1} : Nilai *output* pada *timestep* ke-(t-1)
- b_i : bias untuk *input gate*

Selanjutnya, *input gate* memperbarui informasi dengan membentuk vektor kandidat baru menggunakan fungsi aktivasi tanh yang dirumuskan sebagai berikut.

$$\tilde{c}_t = \tanh(W_{cx} \cdot x_t + W_{ch} \cdot h_{t-1} + b_c) \quad (2.8)$$

dimana:

- \tilde{c}_t : *memory cell* pada *timestep* ke-t
- \tanh : Fungsi tanh
- W_c : Bobot pada *cell state*
- h_{t-1} : Nilai *output* pada *timestep* ke-(t-1)
- b_c : Bias untuk *cell state*

Fungsi tanh atau tangen hiperbolik memiliki rentang nilai dari -1 sampai 1 .
Fungsi tanh dirumuskan sebagai berikut:

$$\tanh x = \frac{\sinh x}{\cosh x}$$

dimana $\sinh x = \frac{e^x - e^{-x}}{2}$ dan $\cosh x = \frac{e^x + e^{-x}}{2}$, sehingga:

$$\tanh x = \frac{\frac{e^x - e^{-x}}{2}}{\frac{e^x + e^{-x}}{2}}$$

$$\tanh x = \frac{e^x - e^{-x}}{2} \cdot \frac{2}{e^x + e^{-x}}$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\tanh x = \frac{e^x(1 - e^{-2x})}{e^x(1 + e^{-2x})}$$

$$\tanh x = \frac{e^x(1 - e^{-2x})}{e^x(1 + e^{-2x})}$$

$$\tanh x = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

$$\tanh x = \frac{2}{1 + e^{-2x}} - \frac{(1 + e^{-2x})}{1 + e^{-2x}}$$

$$\tanh x = \frac{2}{1 + e^{-2x}} - 1$$

Berdasarkan persamaan (2.4), maka diperoleh:

$$\tanh x = 2\sigma(2x) - 1 \quad (2.9)$$

dimana:

σ : fungsi sigmoid.

x : Nilai *input*.

3. Cell state (c_t)

Pada tahap ini, nilai pada *cell state* sebelumnya (c_{t-1}) diperbarui dengan nilai pada *memory cell* yang baru (c_t), melalui persamaan berikut:

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (2.10)$$

dimana:

- c_t : Nilai *cell state* pada *timestep* ke- t
- f_t : *Forget gate* pada *timestep* ke- t
- c_{t-1} : Nilai *cell state* pada *timestep* ke- $(t-1)$
- i_t : Nilai *Input gate* pada *timestep* ke- t
- \tilde{c}_t : *Memory cell* pada *timestep* ke- t

4. *Output gate* (o_t)

Output gate bertujuan untuk mengontrol seberapa banyak keadaan sel saat ini dibuang. Pertama-tama informasi *output* ditentukan oleh lapisan sigmoid, kemudian keadaan sel diproses oleh *tanh* dan dikalikan dengan *output* lapisan sigmoid untuk mendapatkan bagian *output* terakhir, sebagai berikut:

$$o_t = \sigma(W_{ox} \cdot x_t + W_{oh} \cdot h_{t-1} + b_o) \quad (2.11)$$

dimana:

- o_t : *Output gate*
- σ : Fungsi sigmoid
- W_o : Nilai *weight* untuk *output gate*
- h_{t-1} : Nilai *output* sebelum *timestep* ke- t
- x_t : Nilai *input* pada *timestep* ke- t
- b_o : Nilai bias pada *output gate*

Nilai *output* akhir dari sel didefinisikan pada persamaan berikut.

$$h_t = \tanh(c_t) * o_t \quad (2.12)$$

dimana:

- h_t : Nilai *output* pada *timestep* t
- \tanh : Fungsi *tanh*
- c_t : *Cell state*
- o_t : *Output gate*

2.7.2. Bias dan Weight

Hubungan antar *node* diasosiasikan dengan suatu nilai yang disebut dengan bobot atau *weight*. Setiap *node* pasti memiliki *weight*nya masing - masing.

Output merupakan keluaran dari suatu *node*. *Error* merupakan tingkat kesalahan yang terdapat dalam suatu *node* dari proses yang dilakukan. *Weight* merupakan bobot dari *node* tersebut ke *node* yang lain pada *layer* yang berbeda. Bobot awal dalam suatu jaringan syaraf tiruan biasanya diperoleh secara *random* dan sebaiknya diinisialisasi dengan nilai yang relatif kecil. Pada tahap pelatihan, bobot tersebut akan mengalami penyesuaian melalui suatu proses perhitungan matematik agar tercapai nilai bobot yang sesuai.

2.8. Denormalisasi

Proses denormalisasi merupakan proses merubah hasil peramalan yang menggunakan data normalisasi ke dalam bentuk data semula. Proses ini dilakukan untuk menentukan data yang telah diprediksi dan melakukan perbandingan terhadap data aktual guna melihat performa dari model yang telah dibentuk. Jika sebelumnya normalisasi dilakukan dalam interval [0,1], maka denormalisasi dinyatakan dalam persamaan berikut (Wiranda & Sadikin, 2019):

$$y_t = y'(X_{max} - X_{min}) + X_{min} \quad (2.13)$$

dimana:

- y_t : nilai dari data prediksi suhu udara yang telah didenormalisasi
- y' : data prediksi suhu udara yang sudah dinormalisasikan
- X_{min} : nilai minimum data suhu udara
- X_{max} : nilai maksimum data suhu udara

2.9. Evaluasi Model

Evaluasi model merupakan proses pengujian akurasi terhadap performa model yang terbentuk dalam memprediksi suatu data yang telah dilakukan pengujian sebelumnya supaya mengetahui kinerja dari model yang telah dilatih.

Menurut Wiranda & Sadikin (2019), untuk melakukan evaluasi suatu model dapat menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE).

2.9.1. *Root Mean Square Error* (RMSE)

Root Mean Square Error (RMSE) adalah metode yang digunakan dalam mengevaluasi suatu model yang terbentuk untuk menghitung seberapa besar tingkat akurasi. Nilai RMSE diperoleh dari selisih data katual suhu udara secara keseluruhan yang dikurangi data prediksi suhu udara yang dikuadratkan lalu dibagi dengan banyaknya periode peramalan suhu udara dan yang terakhir mengakarkannya. Penentuan nilai RMSE dirumuskan pada persamaan berikut.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n}} \quad (2.14)$$

dimana:

Y_t : Data suhu udara pada periode t.

\hat{Y}_t : Data prediksi suhu udara pada periode t.

n : Banyaknya periode data suhu udara

2.9.2. *Mean Absolute Percentage Error* (MAPE).

Mean Absolute Percentage Error (MAPE) yaitu persentase kesalahan (error) terhadap rata-rata secara absolute (mutlak). Nilai MAPE dapat menghitung

persentase *error* antara nilai data aktual dari data yang diambil dengan nilai hasil prediksinya. MAPE dirumuskan pada persamaan di bawah ini:

$$MAPE = \frac{\sum_{i=1}^n \frac{|\hat{Y}_i - Y_i|}{Y_i}}{n} * 100 \quad (2.15)$$

dimana:

Y_i : Data suhu udara pada periode t

\hat{Y}_i : Data prediksi suhu udara pada periode t

n : Banyaknya periode data suhu udara

Apabila model LSTM memiliki nilai MAPE yang semakin kecil, maka nilai prediksinya mendekati nilai sebenarnya, atau model LSTM dengan nilai MAPE yang terkecil merupakan model LSTM yang terbaik dalam melakukan peramalan suhu udara. Adapun *range* nilai MAPE dalam mengukur kemampuan suatu model LSTM ditampilkan pada Tabel 1 berikut (Hayuningtyas, 2017).

Tabel 1. Kategori *range* nilai MAPE

<i>Range</i> MAPE	Kategori
< 10 %	Kemampuan model peramalan sangat baik
10–20 %	Kemampuan model peramalan baik
20 –50 %	Kemampuan model peramalan layak
>50 %	Kemampuan model peramalan buruk

III. METODOLOGI PENELITIAN

3.1. Waktu dan Tempat Penelitian

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

3.2. Data Penelitian

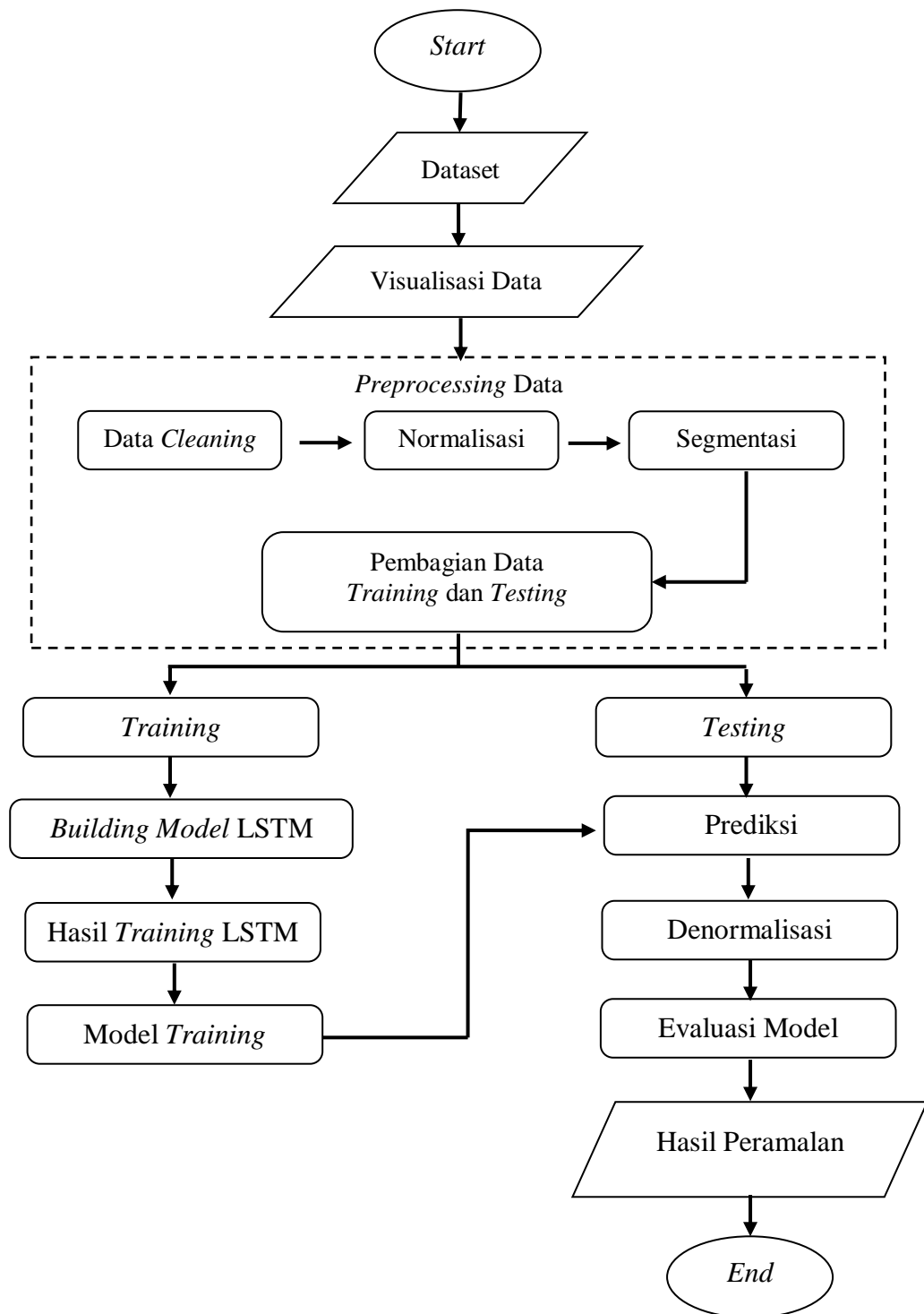
Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari https://dataonline.bmkg.go.id/data_iklim mengenai data histori suhu udara dari tanggal 1 Januari 2018 - 31 Oktober 2022 yang berupa data harian pada Stasiun Klimatologi Lampung. Data berjumlah 1765 berbentuk tabel yang memiliki kolom Tanggal, Suhu udara minimum, suhu udara rata-rata dan suhu udara maksimum. Data Sekunder adalah data yang diperoleh tidak secara langsung dari objek penelitian melainkan data yang sudah jadi yang telah dikumpulkan oleh pihak lain.

3.3. Metode Penelitian

Peramalan suhu udara pada penelitian ini menggunakan metode LSTM dengan memberikan berbagai variasi dari nilai parameter dalam LSTM (*neuron hidden*, *batch size* dan *epoch*). Adapun pengujian metode ini menggunakan bantuan *software* Python yang didukung oleh Google Colab.

Adapun metode atau langkah-langkah dalam pengujian metode ini antara lain sebagai berikut:

1. Tahap pertama dilakukan penginputan data historis suhu udara ke dalam google colab. Proses ini dikenal dengan *Import* dataset.
2. Menampilkan visualisasi data melalui grafik *time series* data untuk melihat aliran data.
3. Lakukan *preprocessing* data yang meliputi data *cleaning* untuk membersihkan adanya *missing value* atau data yang hilang, menormalisasikan data menggunakan metode *Min-Max Scaling*, melakukan segmentasi pada data serta membagi dataset menjadi dua bagian yaitu data untuk *training* dan data untuk *testing*.
4. Membangun model (*building model*) LSTM menggunakan data *training*
5. Menentukan hasil prediksi suhu udara dari model yang terbentuk.
6. Melakukan denormalisasi pada data hasil prediksi.
7. Melakukan evaluasi model menggunakan data *testing* sehingga akan diperoleh nilai RMSE dan MAPE.
8. Membandingkan kinerja model yang terbentuk untuk beberapa variasi jumlah *neuron hidden* (5,25,50,100), *batch Size* (4,16,32,64,128) dan jumlah *epoch* (50,100,150,200) serta berbagai komposisi pembagian data *training* dan data *testing*. Model LSTM dengan nilai MAPE dan RMSE yang terkecil merupakan model LSTM yang terbaik.
9. Menampilkan visualisasi hasil peramalan suhu udara dari model LSTM yang terbaik.
10. Menampilkan hasil peramalan dalam bentuk tabel untuk beberapa waktu kedepan berdasarkan dari model LSTM yang terbaik.
11. Memaparkan kesimpulan.



Gambar 9. Diagram alir metode penelitian

V. KESIMPULAN

Berdasarkan pembahasan, diperoleh kesimpulan dari penelitian ini, diantaranya:

1. Model LSTM yang terbaik pada peramalan suhu udara berdasarkan pengujian komposisi pembagian dataset yaitu model yang terbentuk dari pembagian data training sebesar 80% dan data testing sebesar 20%.
2. Adapun model LSTM yang terbaik untuk peramalan Suhu Udara Minimum dibangun dari 25 *neuron hidden*, 4 *batch size*, dan 50 *epoch*. Sedangkan, model LSTM yang terbaik untuk peramalan Suhu Udara Maksimum dan Suhu Udara Rata-Rata dibangun dari 25 *neuron hidden*, 4 *batch size* dan 50 *epoch*.
3. Nilai RMSE dan MAPE dari model LSTM yang terbaik untuk Suhu Udara Minimum sebesar 0,89 dan 2,59%. Adapun nilai RMSE dan MAPE untuk Suhu Udara Maksimum sebesar 0,85 dan 2,32%. Sedangkan nilai RMSE dan MAPE pada Suhu Udara Rata-Rata sebesar 0,88 dan 2,44%. Karena nilai MAPE kurang dari 10%, maka model LSTM ini sangat baik untuk peramalan suhu udara di Stasiun Klimatologi Lampung.
4. Hasil peramalan Suhu Udara Minimum (X1) dan Suhu Udara Rata-rata (X3) dari Tanggal 01 November 2022 hingga 31 Januari 2023 terus mengalami tren penurunan. Sedangkan, hasil peramalan Suhu Udara Maksimum (X2) mengalami tren peningkatan.

DAFTAR PUSTAKA

- Afriansyah, A.I. 2017. Otomatisasi Pengoperasian Alat Elektronik Berdasarkan Hasil Prediksi Algoritma Long Short Term Memory. *Journal of Information Technology and Computer Engineering*. **4**(2): 83-89.
- 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.
- Aldi, M.W.P., Jondri & Aditsania, A. 2018. Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin. *E-Preceeding of Engineering*. **5**(2): 3548-3555.
- Buduma, N. & Locascio, N. 2017. *Fundamentals of Deep Learning*. O'Reilly Media, United States of America.
- Cahyadi, R., Damayanti, A. & Aryadani, D. 2020. Recurrent Neural Network (RNN) dengan Long Short Term Memory (LSTM) untuk Analisis Sentimen Data Instagram. *Jurnal Informatika dan Komputer (JIKO)*. **5**(1): 1-9.
- Christienova, A. I., Pratiwi, E. W. & Darmawan, G. 2018. Perbandingan Model Peramalan Singular Spectrum Analysis (SSA) dan Fourier Series Analysis (FSA) pada Data Suhu Udara di Surabaya. *BIMIPA*. **94**(1): 94-106.
- Chung, H., & Shin, K. S. 2018. Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*. **10**(10): 3765.
- 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.

- Hochreiter, S. dan Schmidhuber, J. U. 1997 . Long Short-Term Memory. *Neural Comput.*, vol. 9, no. 8, hal. 1735–1780, 1997.
- Karpathy, A., Johnson, J., & Fei-Fei, L. 2015. Visualizing and understanding recurrent networks. *arXiv preprint arXiv*: 1-12.
- Khumaidi, A., Raafi'udin, R. & Solihin, I. P. 2020. Pengujian Algoritma Long Short Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung. *Jurnal Telematika*. **15**(1): 13-18.
- Machmudin, A. & Ulama, B. S. S. 2012. Peramalan Temperatur Udara di Kota Surabaya dengan Menggunakan ARIMA dan Artificial Neural Network. *Jurnal Sains dan Seni ITS*. **1**(1): 118-123.
- Montgomery, D. C., Jennings, L. & Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. John Wiley And Sons Inc, Canada.
- Pratiwi, H. A., Cahyanti, M. & Lamsani, M. 2021. Implementasi Deep Learning Flower Scanner Menggunakan Metode Convolutional Neural Network. *Sebatik*. **25**(1): 124-130
- 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.
- Rizki, M., Basuki, S. & Azhar, Y. 2020. Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory untuk Prediksi Curah Hujan Kota Malang. *Repositor*. **2**(3): 331-338.
- Rusdi, M. 2017. Komparasi Penggunaan Algoritma Support Vector machine dengan particle swarm optimazation dalam memprediksi suhu udara. *Technologia*. **8**(4): 277-284.
- Said, N. Md., Zin, Z. M., Ismail, M. N. & Bakar, T. A. 2021. Univariate Water Consumption Time Series Prediction Using Deep Learning Neural Network (DLLN). *International Journal of Advanced Technology and Engineering Exploration*. **8**(76): 473-483

- Septiyani, M., Helmi & Yudhi. 2019. Penerapan analisis fourier untuk menentukan periode dan prediksi suhu udara (Studi Kasus: Data Suhu Udara Kota Pontianak). *Buletin ilmiah matematika statistika dan terapannya (Bimaster)*. **8**(2): 221-228.
- Supriyadi, E. 2020. Prediksi Parameter Cuaca Menggunakan Deep Learning Long-Short Term Memory (LSTM). *Jurnal Meteorologi dan Geofisika*. **21**(2): 55-67
- Tran, T.D., Tran, V.N. & Kim, J. 2021. Improving the Accuracy of Dam Inflow Predictions Using a Long Short-Term Memory Network Coupled with Wavelet Transform and Predictor Selection. *Mathematics*. **9**: 551.
- 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.
- Yotenka, R. & El Huda, F. F. 2020. Implementasi Long Short Term Memory Pada Harga Saham Perusahaan Perkebunan di Indonesia. *Unisda Journal of Mathematics*. **6**(1): 9-18
- Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. 2017. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*. **11**(2): 68-75.