

**IMPLEMENTASI PERAMALAN SUHU MINIMUM MINGGUAN
MENGUNAKAN METODE *HYBRID SEASONAL AUTOREGRESSIVE
INTEGRATED MOVING AVERAGE (SARIMA)-LONG SHORT TERM
MEMORY (LSTM)***

(Skripsi)

Oleh

ANINDYA DAFA SALSABILA



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

ABSTRACT

IMPLEMENTATION WEEKLY MINIMUM TEMPERATURE FORECASTING USING HYBRID SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA)-LONG SHORT TERM MEMORY (LSTM) METHOD

By

ANINDYA DAFA SALSABILA

Indonesia is a country with a tropical climate that has two seasons in one year, which can influence the weather changes in a region over time, including extreme weather changes. Forecasting is a statistical method that can be used to predict future weather changes. The SARIMA method is a forecasting method that has the advantage of capturing seasonal patterns. However, SARIMA is difficult to use for processing nonlinear components. LSTM is a part of Deep Learning that can process large amounts of data and has three main layers capable of processing fluctuating nonlinear components, namely the input layer, hidden layer, and output layer. Therefore, the hybrid SARIMA-LSTM method is expected to efficiently process data with both linear and nonlinear patterns, including seasonal patterns. The hybrid model used in the research performs well in the data splitting scheme of 80% for training and 20% for testing, with an MSE value of 0.1174, RMSE value of 0.3426, and MAPE value of 0.0104%.

Keywords: Hybrid SARIMA-LSTM, SARIMA, LSTM, Forecasting, Deep Learning, Neural Network, Accuracy.

ABSTRAK

IMPLEMENTASI PERAMALAN SUHU MINIMUM MINGGUAN MENGUNAKAN METODE *HYBRID SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA)-LONG SHORT TERM MEMORY (LSTM)*

Oleh

ANINDYA DAFA SALSABILA

Indonesia merupakan negara yang mempunyai iklim tropis yang memiliki 2 musim dalam 1 tahun yang dapat mempengaruhi perubahan cuaca suatu wilayah setiap waktunya, seperti perubahan cuaca ekstrim. Peramalan merupakan metode statistika yang dapat digunakan untuk memprediksi perubahan cuaca yang akan terjadi di masa depan. Metode SARIMA merupakan metode yang dapat digunakan dalam peramalan dan mempunyai kelebihan dalam menangkap pola musiman. Akan tetapi, SARIMA sulit digunakan untuk memproses komponen nonlinear. LSTM merupakan bagian dari *Deep Learning* yang dapat memproses data berjumlah besar dan mempunyai 3 *layer* utama yang mampu memproses komponen nonlinear data yang berfluktuasi, yaitu *input layer*, *hidden layer* dan *output layer*. Maka dari itu, metode *hybrid* SARIMA-LSTM diharapkan mampu untuk memproses data dengan pola linear dan nonlinear serta *seasonal* secara optimal. Model *hybrid* pada penelitian yang dilakukan memiliki performa yang baik pada skema *splitting data* 80% *training* dan 20% *testing* dengan nilai MSE sebesar 0.1174, RMSE sebesar 0.3426 dan MAPE sebesar 0,0104 %.

Kata Kunci: Hybrid SARIMA-LSTM, SARIMA, LSTM, Peramalan, Deep Learning, Neural Network, Akurasi.

**IMPLEMENTASI PERAMALAN SUHU MINIMUM MINGGUAN
MENGUNAKAN METODE *HYBRID SEASONAL AUTOREGRESSIVE
INTEGRATED MOVING AVERAGE (SARIMA)-LONG SHORT TERM
MEMORY (LSTM)***

Oleh

Anindya Dafa Salsabila

Skripsi

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

Pada

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



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

Judul Skripsi : **IMPLEMENTASI PERAMALAN SUHU
MINIMUM MINGGUAN MENGGUNAKAN
METODE *HYBRID SEASONAL
AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (SARIMA)-LONG SHORT TERM
MEMORY (LSTM)***

Nama Mahasiswa : **Anindya Dafa Salsabila**

Nomor Pokok Mahasiswa : **1917031089**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing


Dian Kurniasari, S.Si., M.Sc.
NIP. 19690305199603 2 001


Prof. Drs. Mustofa, M.A., Ph.D.
NIP. 19570101 198403 1 020

2. Ketua Jurusan Matematika


Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 19740316200501 1 001

MENGESAHKAN

1. Tim Penguji

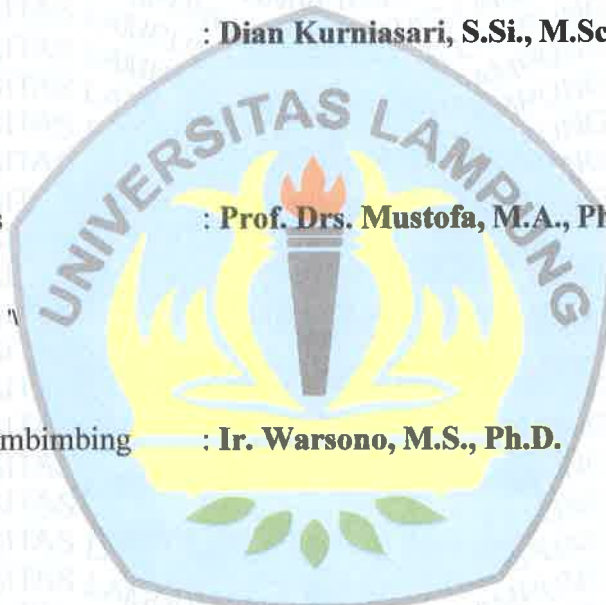
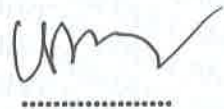
Ketua : **Dian Kurniasari, S.Si., M.Sc**



Sekretaris : **Prof. Drs. Mustofa, M.A., Ph.D.**



Penguji
Bukan Pembimbing : **Ir. Warsono, M.S., Ph.D.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam




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

Tanggal Lulus Ujian Skripsi: **14 Juli 2023**

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini:

Nama : **Anindya Dafa Salsabila**
Nomor Pokok Mahasiswa : **1917031089**
Jurusan : **Matematika**
Judul Skripsi : **IMPLEMENTASI PERAMALAN SUHU
MINIMUM MINGGUAN MENGGUNAKAN
METODE *HYBRID SEASONAL
AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (SARIMA)-LONG SHORT TERM
MEMORY (LSTM)***

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, 14 Juli 2023

Penulis



Anindya Dafa Salsabila

NPM. 1917031089

RIWAYAT HIDUP

Penulis bernama Anindya Dafa Salsabila, anak pertama dari dua bersaudara yang lahir di Kotabumi pada tanggal 9 Juni 2001 dari pasangan Bapak Herman Ramli, SE. dan Ibu Elis Rodiantina

Penulis mengawali pendidikan di Taman Kanak-Kanak An-Nida pada tahun 2006-2007 dan melanjutkan pendidikan Sekolah Dasar di SDN 2 Cilegon pada tahun 2007-2013. Kemudian melanjutkan ke Sekolah Menengah Pertama di SMPN 1 Cilegon pada tahun 2013-2016 dan belajar pada jenjang Sekolah Menengah Atas di SMAN 1 Cilegon pada tahun 2016-2019. Pada tahun 2019 penulis terdaftar sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

Selama aktif menjadi mahasiswa, penulis ikut serta dalam Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) sebagai anggota Bidang Kaderisasi dan Kepemimpinan pada tahun 2020. Pada tahun 2022, penulis melaksanakan Kerja Praktik (KP) di PT. Taspen Kantor Cabang Kota Bandar Lampung pada bidang Kas dan Verifikasi SPJ dan sebagai bentuk pengabdian kepada masyarakat penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Gunung Mekar, Kecamatan Jabung, Kabupaten Lampung Timur.

KATA INSPIRASI

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya”

(Q.S. Al-Baqarah: 286)

“Barangsiapa yang hendak menginginkan dunia, maka hendaklah ia menguasai ilmu. Barangsiapa menginginkan akhirat hendaklah menguasai ilmu dan barangsiapa yang menginginkan keduanya hendaklah ia menguasai ilmu”

(HR Ahmad)

“Balas dendam terbaik adalah menjadikan dirimu lebih baik”

(Ali Bin Abi Thalib)

*“Kejarlah ilmu, gapai cita-citamu dan bungkam mereka atas pencapaianmu.
Ketekunanmu akan membuktikan.*

(Penulis)

PERSEMBAHAN

Dengan mengucapkan Alahmdulillahirabbil'alamin, segala puji dan syukur kepada Allah SWT atas nikmatnya sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya.

Diri Sendiri

Terima kasih atas segala ketekunan dan tetap berusaha dalam segala keadaan untuk mengembangkan diri menjadi pribadi yang lebih baik.

Orang Tersayang

Terima kasih kepada kalian yang telah tulus dan memberikan doa, dukungan, motivasi, waktu, pengorbanan, cinta dan kasih sayang kepada saya. Terima kasih atas segala ridho dan doa yang senantiasa diberikan dalam setiap langkah yang saya hadapi untuk tetap mampu berproses dan menggapai cita-cita.

Dosen Pembimbing dan Pembahas

Terimakasih kepada dosen pembimbing dan pembahas yang dengan tulus memberikan ilmu, dukungan, motivasi dan arahan yang berharga kepada penulis.

Teman-teman yang telah berjuang bersama dan kebersamai serta mendoakan saya dalam setiap perjalanan saya.

Almamater Tercinta Universitas Lampung

SANWACANA

Puji syukur saya ucapkan kehadirat Allah SWT, atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Implementasi Peramalan Suhu Minimum Mingguan Menggunakan Metode *Hybrid Seasonal Autoregressive Integrated Moving Average (SARIMA)-Long Short Term Memory (LSTM)*”. Dalam kesempatan ini, penulis berterima kasih kepada:

1. Ibu Dian Kurniasari, S.Si., M.Sc. selaku Dosen Pembimbing I yang senantiasa memberikan arahan, bantuan, motivasi dan saran kepada penulis dalam penyusunan skripsi ini.
2. Bapak Prof. Drs. Mustofa, M.A., Ph.D., selaku Dosen Pembimbing II yang telah memberikan bimbingan serta saran yang membantu kepada penulis dalam proses penyelesaian skripsi ini.
3. Bapak Ir. Warsono, M.S., Ph.D., selaku Dosen Pembahas yang telah memberikan masukan, kritik dan saran serta evaluasi kepada penulis dalam penyelesaian skripsi ini.
4. Ibu Dr. Fitriani, S.Si., M.Sc., selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan dan saran kepada penulis selama perkuliahan.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Dr. Eng. Heri Satria, S.Si., M.Si., selaku dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen, staf, dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Ayah, ibu, adik dan seluruh keluarga besar yang senantiasa memberikan kasih sayang, dukungan, nasihat, motivasi doa, serta segala kebutuhan

kepada penulis demi kemudahan dan kelancaran untuk menyelesaikan penulisan ini.

9. Fiqih yang telah mendukung penulis untuk tetap berproses dalam menyelesaikan perkuliahan.
10. Rizke, Zida dan Berlian yang selalu menyemangati dan memotivasi serta memberikan kenangan indah selama masa perkuliahan.
11. Teman-teman Matematika 2019 yang telah menjadi rekan seperjuangan selama perkuliahan.
12. Seluruh pihak terkait yang membantu dalam menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari laporan ini masih jauh dari kata sempurna dan masih terdapat banyak kekurangan baik, namun penulis berharap semoga skripsi ini dapat berguna dan bermanfaat bagi pembaca yang membutuhkan. Oleh sebab itu, saran dan kritikan yang membangun senantiasa penulis harapkan demi menyempurnakan skripsi ini.

Bandar Lampung, 17 Juli 2023

Penulis

Anindya Dafa Salsabila

DAFTAR ISI

	Halaman
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	1
I. PENDAHULUAN.....	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian.....	5
1.3 Manfaat Penelitian.....	5
II. TINJAUAN PUSTAKA	6
2.1 Peramalan	6
2.2 <i>Time Series</i>	7
2.3 Stasioneritas.....	11
2.4 Diagnostik Model	12
2.4.1 Residual Bersifat Acak.....	12
2.4.2 Residual Berdistribusi Normal.....	13
2.5 <i>Data Mining</i>	14
2.6 <i>Machine Learning</i>	15
2.7 <i>Deep Learning</i>	16
2.7.1 <i>Recurrent Neural Networks</i>	17
2.7.2 <i>Long Short Term Memory</i>	18
2.7.3 Fungsi Aktivasi	22
2.7.4 <i>Hyperparameter</i>	24
2.8 <i>Scaling Data</i>	25
2.9 Evaluasi Model.....	26
2.10 <i>Hybrid SARIMA-LSTM</i>	28

III. METODOLOGI PENELITIAN.....	29
3.1 Waktu dan Tempat Penelitian	29
3.2 Spesifikasi Perangkat Penelitian	29
3.3 Data Penelitian	29
3.4 Metode Penelitian.....	31
IV. HASIL DAN PEMBAHASAN.....	34
4.1 <i>Input Data</i>	34
4.2 <i>Preprocessing Data</i>	35
4.3 Visualisasi Data	36
4.4 Prediksi Model SARIMA.....	37
4.4.1 Stasioneritas	37
4.4.2 Identifikasi Model SARIMA.....	39
4.4.3 Pemilihan Model SARIMA Terbaik	41
4.4.4 Prediksi Model SARIMA.....	43
4.4.1 Residual Model SARIMA.....	45
4.5 Prediksi Model <i>Hybrid</i> SARIMA-LSTM.....	46
4.5.1 <i>Preprocessing Data II</i>	47
4.5.2 Membangun Model Prediksi SARIMA dengan LSTM	47
4.5.3 Membangun Model Residual SARIMA dengan LSTM	49
4.5.4 Prediksi Model <i>Hybrid</i> SARIMA-LSTM	51
4.5.5 Peramalan Model <i>Hybrid</i> SARIMA-LSTM.....	52
V. KESIMPULAN DAN SARAN.....	54
5.1 Kesimpulan.....	54
5.2 Saran.....	55
DAFTAR PUSTAKA	56
LAMPIRAN	

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Kriteria nilai MAPE	27
Tabel 2. Data Penelitian	30
Tabel 3. Data <i>Input</i>	34
Tabel 4. Suhu Minimum Mingguan	35
Tabel 5. Uji ADF Data <i>Input</i>	38
Tabel 6. Uji ADF Musiman	38
Tabel 7. Identifikasi Model SARIMA Sementara	40
Tabel 8. Pendugaan Parameter Model	42
Tabel 9. Data Prediksi dan <i>Invers Differencing</i>	44
Tabel 10. Akurasi Prediksi Model SARIMA	45
Tabel 11. Skema Pembagian <i>Dataset</i>	47

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Struktur RNN	18
Gambar 2. Arsitektur LSTM	19
Gambar 3. Grafik Fungsi Sigmoid (Logistik)	23
Gambar 4. Grafik Fungsi Tanh	23
Gambar 5. Diagram Alir Metode Penelitian	33
Gambar 6. Plot Data Suhu Minimum.....	36
Gambar 7. Plot Dekomposisi	36
Gambar 8. Plot ACF dan PACF	39
Gambar 9. <i>Summary</i> Model SARIMA	43
Gambar 10. Plot Data Aktual dan Prediksi Model SARIMA	44
Gambar 11. Plot Data Terkini dan Peramalan Model SARIMA.....	45
Gambar 12. Plot Residual Model SARIMA.....	46
Gambar 13. Proses <i>Hypertunning</i> Model Prediksi SARIMA dengan LSTM	48
Gambar 14. <i>Best Model</i> Prediksi SARIMA dengan LSTM	49
Gambar 15. Proses <i>Hypertunning</i> Model Residual SARIMA dengan LSTM	50
Gambar 16. <i>Best Model</i> Residual SARIMA dengan LSTM	50
Gambar 17. Visualisasi dan Evaluasi Model <i>Hybrid</i> SARIMA-LSTM.....	51
Gambar 18. Visualisasi Peramalan Model <i>Hybrid</i> SARIMA-LSTM	52
Gambar 19. Visualisasi Peramalan Model <i>Hybrid</i> SARIMA-LSTM (80% <i>Training</i> dan 20% <i>Testing</i>).....	53

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Cuaca merupakan suatu keadaan udara pada wilayah tertentu yang relatif sempit dalam jangka waktu yang relatif singkat (BMKG, 2021). Sedangkan iklim merupakan keadaan cuaca minimum pada wilayah tertentu yang luas dalam waktu yang relatif lama (BMKG, 2021). Indonesia merupakan negara yang mempunyai iklim tropis dikarenakan terletak disekitar garis khatulistiwa sehingga memiliki 2 musim dalam 1 tahun, yaitu musim penghujan dan musim kemarau. Perubahan iklim mempengaruhi perubahan cuaca suatu wilayah setiap waktunya, seperti cuaca ekstrim, kenaikan temperatur udara, perubahan pola curah hujan dan kenaikan muka air laut dapat berdampak pada kesehatan (Bappenas, 2010). Dalam bidang kesehatan, penyakit malaria, demam berdarah (DBD) dan diare memiliki tingkat kejadian sangat tinggi dan penyebarannya yang luas di Indonesia akibat perubahan iklim.

Peramalan atau *forecasting* merupakan metode statistika yang dapat digunakan untuk memprediksi keadaan cuaca yang akan terjadi di masa depan. Prediksi tersebut dapat digunakan sebagai salah satu acuan bagi pemerintah dalam menyusun strategi penanggulangan risiko dan dapat digunakan masyarakat luas dalam menghadapi perubahan cuaca ekstrim. Penggunaan metode peramalan disesuaikan dengan mengidentifikasi bentuk data (Montgomery, dkk., 2015). Pemilihan metode peramalan disesuaikan dengan pola data yang terbentuk, faktor yang mempengaruhi hasil peramalan (Aksan & Khalilah, 2020). Peramalan dengan analisis runtun waktu tertentu merupakan peramalan deret waktu (*time*

series). Data *time series* memiliki empat bentuk pola data, yaitu pola konstan (*horizontal*), pola *trend*, pola siklus (*cycle*) dan pola musiman (*seasonal*). Pada dasarnya, peramalan *time series* merupakan nilai di masa depan yang berasal dari nilai masa lampau (Aksan & Khalilah, 2020).

Pola musiman merupakan salah satu bentuk pola data deret waktu yang memiliki pengulangan pola setiap beberapa periode waktu tertentu, seperti satu bulan, triwulan, satu tahun, dan sebagainya (Petropoulos, dkk., 2022). Salah satu metode peramalan *time series* yang dapat digunakan untuk meramalkan keadaan cuaca adalah metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Metode SARIMA merupakan pengembangan dari metode ARIMA yang memiliki pola musiman (Chen, dkk., 2018). Metode SARIMA merupakan pengembangan dari metode ARIMA yang dapat menganalisis plot data musiman, maka dasar ilmu metode SARIMA sama seperti metode ARIMA. Model ARIMA dapat diperluas ke model SARIMA dimana dapat ditulis dengan $ARIMA(p,d,q)(P,D,Q)_s$ (Petropoulos, dkk., 2022). Keterbatasan model SARIMA yaitu memiliki asumsi linearitas dan residual yang dihasilkan masih terdapat unsur nonlinear (Wu, dkk., 2021). Metode SARIMA memiliki akurasi yang tinggi untuk peramalan jangka waktu yang pendek, tetapi akurasinya akan menurun pada saat digunakan untuk peramalan jangka panjang. Salah satu metode yang dapat digunakan dalam penyelesaian permasalahan tersebut adalah dengan menggunakan metode *Long Short Term Memory* (LSTM).

Model SARIMA dan LSTM memiliki kelebihan dan kekurangan masing-masing. Sebuah model gabungan (*hybrid*) dirancang untuk saling menutupi kekurangan dari tiap model dan memanfaatkan kelebihan masing-masing model sehingga dapat meningkatkan akurasi dari hasil peramalannya (Rowan, dkk., 2022). *Long Short Term Memory* (LSTM) merupakan jenis arsitektur pengembangan dari *Recurrent Neural Network* (RNN), dengan menambahkan *memory cell* sehingga memiliki kelebihan untuk dapat menyimpan informasi dalam jangka panjang. Model LSTM mampu dalam memproses data bersifat linear maupun nonlinear, namun dalam memproses data akan membutuhkan waktu yang lama. Metode

LSTM cocok digunakan untuk data *time series* dalam mengklasifikasi, memproses dan membuat prediksi dalam menangkap suatu hubungan fungsional yang tidak diketahui (Yadav, dkk., 2020). Metode LSTM dapat mengatasi *vanishing gradient* yang menjadi permasalahan dari RNN dalam menangkap *long term dependencies* yang dapat mempengaruhi akurasi dari suatu prediksi (Wiranda & Sadikin, 2019). Model LSTM memasukkan nilai residual yang dihasilkan oleh model SARIMA yang tetap mempertahankan komponen nonlinearitas ke dalam lapisan *input* LSTM (Wu, dkk., 2021).

Penelitian oleh Sirisha, dkk. (2022), melakukan analisis keuntungan untuk memprediksi keuntungan dan kerugian penjualan di masa depan dari suatu produk dengan menggunakan metode ARIMA, SARIMA dan LSTM. Dari membandingkan metode tersebut diperoleh hasil akurasi tertinggi dengan menggunakan metode LSTM dengan nilai RMSE sebesar 3,9172.

Penelitian tentang *hybrid* SARIMA-LSTM sudah pernah dilakukan sebelumnya. Penelitian yang dilakukan oleh Padhilha, dkk. (2022), melakukan penelitian terkait peningkatan energi terbarukan dengan menguji pada 5 data *time series* yang berbeda dengan membandingkan 9 metode yaitu, metode SARIMA, RNN, LSTM, GRU, TRANSFORMER, SARIMA+RNN, SARIMA+LSTM, SARIMA+GRU dan H-TRANSFORMER. Hasil yang diperoleh yaitu dengan menggunakan metode *hybrid* memiliki akurasi yang lebih baik dibandingkan metode tunggal pada tiap pengujian dengan data yang berbeda, dimana pada pengujian data Wind1 metode *hybrid* SARIMA+LSTM merupakan metode terbaik dalam peramalan dengan nilai RMSE sebesar 0,611.

Penelitian yang dilakukan oleh Wu, dkk. (2021), melakukan peramalan tentang kedatangan turis harian dari masing-masing 6 negara menggunakan metode *hybrid* SARIMA-LSTM. Hasil yang diperoleh yaitu dibandingkan metode ARIMA, SARIMA, LSTM, *Naive*, *Seasonal Naive*, SES dan *Holt Winter's*,

performa dari metode *hybrid* SARIMA-LSTM lebih baik pada tiap 6 negara berbeda dengan menghasilkan MAPE terkecil sebesar 0,1849.

Penelitian oleh Sun, dkk. (2020), yaitu memprediksi kenaikan permukaan air laut minimum di Cina dan sekitarnya menggunakan metode SARIMA, LSTM, *hybrid* SARIMA-LSTM dan *hybrid* Metode Kuadrat Terkecil (LS)-Fungsi Basis Radial (RBF). Penelitian ini memberikan hasil RMSE terbaik menggunakan metode *hybrid* SARIMA-LSTM sebesar 1,155.

Penelitian oleh Tahyudin, dkk. (2022), melakukan peramalan tentang jumlah pasien COVID-19 dengan metode SARIMA, LSTM dan *hybrid* SARIMA-LSTM. Hasil yang di peroleh yaitu performa dari metode *hybrid* SARIMA-LSTM lebih baik dibandingkan dengan metode SARIMA maupun metode LSTM, dengan menghasilkan RMSE sebesar 0,33905765.

Penelitian oleh Ding, X. W. (2022), terkait peramalan omset perdagangan bilateral antara China dan Rusia di periode pasca pandemi pada kegiatan ekspor-impor. Penelitian ini membandingkan beberapa model kombinasi menggunakan kombinasi metode linear dan nonlinear, yaitu model SARIMA, LSTM, SVR, SARIMA-LSTM, LSTM-SARIMA, SARIMA-SVR dan SVR-SARIMA di dapatkan hasil akurasi ramalan terbaik menggunakan metode *hybrid* SARIMA-LSTM dengan nilai RMSE sebesar 0,8618.

Berdasarkan penjelasan diatas, penulis tertarik untuk melakukan peramalan mengenai keadaan cuaca di Kota Bandar Lampung menggunakan metode *hybrid* SARIMA-LSTM. Pemilihan Kota Bandar Lampung dipilih karena penulis yang saat ini menetap di Kota Bandar Lampung, sehingga dapat memudahkan dalam pengumpulan data yang diperlukan.

1.2 Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Membangun model dengan metode *hybrid* SARIMA-LSTM.
2. Memprediksi model *hybrid* SARIMA-LSTM.
3. Mengevaluasi model dari metode *hybrid* SARIMA-LSTM dari model terbaik dengan melihat nilai MAPE, RMSE dan MAE.
4. Meramalkan keadaan cuaca di Kota Bandar Lampung dengan model terbaik..

1.3 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

1. Menambah wawasan penulis terhadap pengaplikasian metode *hybrid* SARIMA-LSTM.
2. Dapat menjadi referensi bagi peneliti selanjutnya untuk jenis data lain dengan menggunakan metode *hybrid* SARIMA-LSTM.

II. TINJAUAN PUSTAKA

2.1 Peramalan

Peramalan merupakan metode statistika yang dapat digunakan dalam prediksi beberapa peristiwa di masa depan yang mencakup banyak bidang, seperti masalah bisnis dan industri, pemerintah, ekonomi, ilmu lingkungan, kedokteran, ilmu sosial dan keuangan (Montgomery, dkk., 2015). Peramalan sering diklasifikasikan sebagai perkiraan masalah jangka pendek, jangka menengah dan jangka panjang. Peramalan jangka pendek mencakup prediksi peristiwa di masa depan dalam beberapa periode waktu (hari, minggu, bulan). Peramalan jangka menengah diperpanjang dari satu hingga dua tahun ke masa depan, sedangkan masalah jangka panjang dapat melampaui beberapa tahun kedepan (Montgomery, dkk., 2015).

Teknik peramalan dibagi menjadi dua jenis, yaitu metode kualitatif dan metode kuantitatif (Montgomery, dkk., 2015), yaitu:

1. Peramalan kualitatif merupakan metode yang didasarkan atas pengambilan data kualitatif pada masa lalu. Hasil peramalan kualitatif didasarkan pada pengamatan di masa lalu yang di gabungkan dengan pemikiran penyusunnya, seperti intuisi pengambilan keputusan, emosi, pengalaman pribadi dan sistem nilai.
2. Peramalan kuantitatif merupakan metode yang didasari atas pengambilan data kuantitatif pada masa lalu yang berasal dari pengamatan nilai-nilai sebelumnya. Hasil peramalan kuantitatif sangat dipengaruhi pada metode

yang digunakan dalam peramalan, sehingga setiap mode akan memiliki hasil peramalan yang berbeda-beda.

Terdapat dua jenis model peramalan pada peramalan kuantitatif yaitu model deret waktu (*time series*) dan model regresi (*regression*). Pada penelitian ini menggunakan data historis dengan interval waktu harian, sehingga model penelitian ini menggunakan model deret waktu (*time series*).

2.2 *Time Series*

Deret waktu merupakan serangkaian pengamatan terhadap suatu variabel yang diamati dari waktu ke waktu dan dicatat berdasarkan periode waktu tertentu dalam interval waktu yang tetap (Wei, 2006). Analisis deret waktu atau *time series* adalah suatu metode analisis statistika yang digunakan untuk peramalan dengan data *time series* atau data runtun waktu berkala. Pemilihan metode peramalan disesuaikan dengan pola data yang terbentuk, faktor yang mempengaruhi hasil peramalan. Dalam analisis deret waktu terdapat 4 pola data yang dapat digunakan sebagai pertimbangan pemilihan metode yang tepat, yaitu (Makridakis, dkk., 1995):

1. Pola konstan

Pola data konstan terjadi apabila data pengamatan berfluktuasi di sekitar suatu nilai konstan atau *mean* yang membentuk garis lurus, dapat disebut juga sebagai data stasioner. Misalnya penjualan suatu produk yang penjualannya tidak meningkat atau menurun dalam periode waktu tertentu.

2. Pola *trend*

Pola data *trend* terjadi apabila data pengamatan mengalami kenaikan atau penurunan dalam periode waktu tertentu berjangka panjang, sehingga data dengan pola *trend* disebut juga sebagai data nonstasioner.

3. Pola musiman

Pola data musiman terjadi apabila data pengamatan dipengaruhi oleh musiman. Pola musiman memiliki pola berulang setelah beberapa periode waktu tertentu, seperti satu tahun, triwulan, satu bulan, dan sebagainya (Petropoulos, dkk., 2022). Misalnya peningkatan jumlah wisatawan saat libur sekolah.

4. Pola siklis

Pola data siklis terjadi apabila data pengamatan dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang terjadi pada siklus bisnis.

Metode *time series* dapat dibedakan menjadi beberapa teknik, diantaranya adalah metode Box-Jenkins atau ARIMA. Bentuk umum dari model AR, MA, ARMA, ARIMA dan SARIMA yaitu:

1. Model *Autoregressive* (AR)

Model AR merupakan model stasioner dari *time series* yang mengasumsikan bahwa suatu peristiwa saat ini dipengaruhi oleh kejadian yang sama pada periode sebelumnya. Model $AR(p)$ secara umum ditulis sebagai (Montgomery, dkk., 2015):

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t \quad (2.1)$$

dimana:

Z_t = nilai pengamatan pada waktu ke- t , $t = 1, 2, \dots, n$

Z_{t-1}, \dots, Z_{t-p} = nilai masa lalu pada waktu ke- $t-1, \dots, t-p$

ϕ_i = parameter AR tingkat- i , $i = 1, 2, \dots, p$

ε_t = nilai *error* pada waktu ke- t

p = *orde* AR non-musiman

2. Model *Moving Average* (MA)

Model MA mengasumsikan bahwa prediksi suatu nilai saat ini dipengaruhi oleh *error* periode sebelumnya. Model $MA(q)$ secara umum ditulis sebagai (Montgomery, dkk., 2015):

$$Z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (2.2)$$

dimana:

Z_t	= nilai pengamatan pada waktu ke- t , $t = 1, 2, \dots, n$
Z_{t-1}, \dots, Z_{t-q}	= nilai masa lalu pada waktu ke- $t-1, \dots, t-q$
θ_j	= parameter MA tingkat j , $j = 1, 2, \dots, q$
ε_{t-j}	= nilai <i>error</i> pada waktu ke $t - j$, $j = 1, 2, \dots, q$
ε_t	= nilai <i>error</i> pada waktu ke- t
q	= <i>orde</i> MA non-musiman

3. Model *Autoregressive Moving Average* (ARMA)

Model ARMA menggabungkan metode AR dan MA yang mengasumsikan bahwa peristiwa saat ini dipengaruhi oleh kejadian yang sama dan nilai *error* pada periode sebelumnya. Model ARMA(p, q) secara umum ditulis sebagai (Montgomery, dkk., 2015):

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (2.3)$$

dimana:

Z_t	= nilai variabel pada waktu ke- t , $t = 1, 2, \dots, n$
Z_{t-1}, \dots, Z_{t-p}	= nilai masa lalu pada waktu ke- $t-1, t-2, \dots, t-p$
ϕ_i	= parameter AR tingkat- i , $i = 1, 2, \dots, p$
θ_j	= parameter MA tingkat j , $j = 1, 2, \dots, q$
ε_{t-q}	= nilai <i>error</i> pada waktu ke $t - q$
p	= <i>orde</i> AR non-musiman
q	= <i>orde</i> MA non-musiman

4. Model *Autoregressive Integrated Moving Average* (ARIMA)

Model *Autoregressive Integrated Moving Average* (ARIMA) merupakan model ARMA nonstasioner yang telah dilakukan *differencing* hingga menjadi model stasioner. Adapun model ARIMA(p, d, q), yaitu sebagai berikut (Montgomery, dkk., 2015):

$$(1 - \phi_1 b)(1 - B)Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_t \quad (2.4)$$

atau

$$Z_t - Z_{t-1} = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_t$$

$$Z_t = Z_{t-1} + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_t \quad (2.5)$$

dimana:

$$\phi_p(B) = 1 - \phi_1(B) - \phi_2 B^2 - \dots - \phi_p B^p \quad \text{adalah operator AR} \quad (2.6)$$

$$\theta_q(B) = 1 - \theta_1(B) - \theta_2 B^2 - \dots - \theta_q B^q \quad \text{adalah operator MA} \quad (2.7)$$

B = operator *backshif*

$(1 - B)^d Z_t$ = deret waktu yang stasioner pada *differencing* ke- d

5. Model *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

Metode SARIMA merupakan pengembangan dari metode ARIMA yang dapat menganalisis masalah pada data *time series* yang memiliki pola musiman (Chen, dkk., 2018). Model SARIMA terdiri dari dua bagian, yaitu bagian tidak musiman dan bagian musiman. Metode SARIMA memiliki tingkat akurasi yang tinggi dalam prediksi jangka pendek. Model ARIMA dapat diperluas ke model SARIMA yang dapat ditulis dengan $ARIMA(p,d,q)(P,D,Q)_s$ (Petropoulos, dkk., 2022). Model SARIMA dapat dinotasikan sebagai:

$$ARIMA(p, d, q)(P, D, Q)_s \quad (2.8)$$

dimana:

(p, d, q) = bagian tidak musiman model

(P, D, Q) = bagian musiman model

P = orde musiman AR

Q = orde musiman MA

D = banyaknya *differencing* musiman

S = jumlah periode tiap musim

Persamaan umum SARIMA sebagai berikut:

$$\begin{aligned} \phi_p(B)(1 - B)^d \Phi_P(B^S)(1 - B^S)^D x Z_t &= \theta_q(B)\theta_Q(B^S)\varepsilon_t \\ (1 - \phi_p B)(1 - B)(1 - \Phi_P B^S)(1 - B^S)Z_t &= (1 - \theta_q B)(1 - \theta_Q)\varepsilon_t \\ Z_t = Z_{t-1} + \phi_1 Z_{t-1} - \phi_1 Z_{t-2} + \dots + \phi_1 Z_{t-S} - \phi_1 Z_{t-S-1} + \dots - \phi_P Z_{t-2S} + \\ &\quad \Phi_P Z_{t-2S-1} - \theta_q \varepsilon_{t-q} + \dots - \theta_Q \varepsilon_{t-S} + \varepsilon_t \end{aligned} \quad (2.9)$$

dimana:

Z_t = pengamatan pada waktu ke- t

$\phi_p(B)$ = operator AR

$\Phi_p(B^S)$ = operator AR musiman

θ_q = operator MA

$\Theta_p(B^S)$ = operator MA musiman

$(1 - B)^d$ = orde *diff* non-musiman

$(1 - B^S)^D$ = orde *diff* musiman

ε_t = nilai *error* pada periode t

2.3 Stasioneritas

Stasioner menunjukkan bahwa tidak adanya perubahan yang drastis pada data, yang dapat diidentifikasi dari fluktuasi bentuk sebaran data disekitar nilai minimum yang konstan, tidak bergantung pada waktu dan varians dari fluktuasi tersebut (Makridakis, dkk, 1995). Apabila sebelum dilakukannya analisis data diketahui tidak stasioner maka data perlu di stasionerkan terlebih dahulu dengan metode yang tepat. Suatu data *time series* dapat dikatakan stasioner dalam minimum jika minimumnya tetap, yaitu tidak memiliki pola *trend* naik maupun turun. Data yang tidak stasioner dalam minimum dapat dilakukan proses *differencing* (pembedahan). *Differencing* merupakan proses mencari selisih antara data satu periode (Z_t) dengan periode sebelumnya (Z_{t-1}). Proses *differencing* (ΔZ_t) pada orde ke- d dapat dilakukan hingga data stasioner (Makridakis, dkk., 1995). Persamaan *differencing* orde pertama dapat ditulis sebagai:

$$\begin{aligned}\Delta Z_t &= Z_t - Z_{t-1} \\ &= Z_t - BZ_t \\ &= (1 - B)Z_t\end{aligned}\tag{2.10}$$

dengan Z_t adalah data asli setelah dilakukan *differencing* orde pertama ($d = 1$) pada Persamaan (2.10) dinyatakan $(1 - B)$ dan notasi B adalah operator shift mundur (*backward shift*). Apabila setelah dilakukan stasioner orde pertama data masih belum stasioner, maka akan dilakukan kembali *differencing* orde kedua ($d = 2$). Persamaan *differencing* orde kedua dapat ditulis sebagai:

$$\begin{aligned}
 \Delta^2 Z_t &= \Delta(\Delta Z_t) \\
 &= \Delta(Z_t - \Delta Z_{t-1}) \\
 &= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) \\
 &= Z_t - 2Z_{t-1} + Z_{t-2} \\
 &= Z_t - 2BZ_t + B^2Z_t \\
 &= (1 - B)^2 Z_t
 \end{aligned} \tag{2.11}$$

Differencing orde kedua ($d = 2$) pada Persamaan (2.11) dinyatakan $(1 - B)^2$. Apabila terdapat *differencing* orde ke- d untuk mencapai stasioner, maka persamaan *differencing* orde- d dapat dinyatakan sebagai:

$$\Delta^d Z_t = (1 - B)^d Z_t, \quad d \geq 1 \tag{2.12}$$

2.4 Diagnostik Model

Pemeriksaan diagnostik model terdiri dari uji *white noise* dan uji normalitas. Suatu model yang baik akan bersifat *white noise*, yaitu tidak adanya autokorelasi residual dan berdistribusi normal.

2.4.1 Residual Bersifat Acak

White noise merupakan proses dimana residual tidak memiliki korelasi dan dapat di deteksi menggunakan uji autokorelasi pada analisis residualnya (Wei, 2006). Uji *white noise* model dikatakan baik apabila residual bersifat acak yang

menunjukkan tidak terdapat autokorelasi dan residual tidak membentuk pola tertentu. Berikut merupakan uji statistik menggunakan uji *Ljung-Box* yang dapat digunakan dalam mendeteksi adanya autokorelasi dari suatu model (Montgomery, dkk., 2015):

Hipotesis:

$H_0: r_1 = r_2 = \dots = r_k = 0$ (Residual tidak berautokorelasi)

$H_1: r_i \neq r_j$ (Residual berautokorelasi)

Taraf signifikan (α) yang digunakan sebesar $5\% = 0,05$

Statistik uji yang digunakan:

$$LB = n(n + 2) \sum_{k=1}^m \frac{r_k}{n - k} \quad (2.13)$$

dimana:

n = banyaknya data pengamatan

k = nilai *lag*

m = nilai maksimum *lag* yang di uji

r_k = nilai autokorelasi pada *lag* ke- k

Kriteria uji:

Terima H_0 yaitu residual tidak berautokorelasi, jika $p - value > 0,05$ atau nilai $LB < X_{(\alpha, df)}^2$ dengan derajat bebas $(df) = k - p$ dan p adalah banyaknya parameter.

2.4.2 Residual Berdistribusi Normal

Uji residual berdistribusi normal atau uji normalitas digunakan untuk memeriksa apakah suatu residual berdistribusi secara normal atau tidak berdasarkan data yang diperoleh. Pengujian normalitas residual dapat dilakukan dengan statistik uji *Kolmogorov-Smirnov* sebagai berikut (Montgomery, dkk., 2015).

Hipotesis:

H_0 : residual berdistribusi secara normal

H_1 : residual tidak berdistribusi secara normal

Taraf signifikan (α) yang digunakan sebesar $5\% = 0,05$

Statistik uji yang digunakan:

$$D = KS = \max|F_0(X) - S_n(X)| \quad (2.14)$$

dimana:

$F_0(X)$ = fungsi distribusi kummulatif perbanding

$S_n(X)$ = fungsi distribusi kumulatif observasi

Kriiteria uji:

Terima H_0 yaitu residual berdistribusi normal, jika $p - value > 0,05$ atau

$D_{hitung} < D_{(\alpha,n)}$.

2.5 Data Mining

Data mining merupakan proses menemukan informasi baru dengan mencari pola dan *trend* yang berguna dalam kumpulan data yang sangat besar (Larose & Larose, 2014). *Data mining* dapat dikatakan sebagai serangkaian proses gabungan dari teknik statistik, matematika, kecerdasan buatan dan *machine learning* dalam menggali nilai tambah berupa pengetahuan yang tidak diketahui dan meringkas data sehingga mudah dipahami dan dimengerti dari suatu kumpulan data pengamatan berukuran besar (Hand, dkk., 2001). *Data mining* disebut juga dengan istilah *knowledge discovery in database* (KDD). Teknik *data mining* berasal dari campuran teknik analisis data dengan algoritma-algoritma dalam memproses data yang besar untuk menemukan pola-pola baru dan informasi yang berguna. *Data mining* dapat dilakukan untuk semua jenis data sesuai kebutuhan yang ingin dicapai. *Data mining* menggabungkan berbagai bidang ilmu lain, seperti statistika, sistem basis data, *data warehouse*, *machince learning* dan *information retrieval* (Hand, dkk., 2012). Beberapa metode *data*

mining yang paling sering digunakan, antara lain estimasi, prediksi, klasifikasi dan *clustering* (Larose & Larose, 2014).

2.6 *Machine Learning*

Machine Learning merupakan bagian dari cabang ilmu kecerdasan buatan (*Artificial Intelligence*) yang erat kaitannya dengan bagaimana membangun program komputer dalam meningkatkan kecerdasannya berdasarkan data untuk pengenalan pola, prediksi dan klasifikasi (Muntiari & Hanif, 2022). *Machine Learning* mampu dalam menangani data berdimensi tinggi dan memetakannya menjadi beberapa kelas dengan karakteristik yang kompleks. Tujuan dari *machine learning* pada umumnya ada dua, yaitu memprediksi suatu peristiwa di masa yang akan datang (*unobserved event*) dan memperoleh ilmu pengetahuan (*knowledge discovery/discovering unknown structure*). Tujuan tersebut akan tercapai melalui *dataset* atau kumpulan data/sampel dalam statistik, dilanjutkan dengan membangun model untuk menggeneralisasi suatu aturan maupun pola data sehingga mendapatkan suatu informasi atau mengambil keputusan (Hand, 2012).

Machine Learning terbagi menjadi 3 teknik pembelajaran, yaitu:

1. *Supervised Learning*

Supervised Learning merupakan pembelajaran terarah atau terawasi yang ditandai dengan adanya *class/label/target* untuk mengklasifikasikan kelas yang tidak dikenal pada himpunan data. *Supervised Learning* adalah sebuah pendekatan dimana sudah tersedianya data yang dilatih dan memiliki variabel yang ditargetkan sehingga tujuan dari pendekatan ini yaitu mengelompokkan suatu data ke data yang sudah ada (Pradnyana & Agustini, 2022). *Supervised Learning* secara umum bertujuan untuk melakukan prediksi dan klasifikasi (*classification*). Beberapa algoritma yang termasuk *Supervised Learning*, yaitu *Backpropagation*, *Random Forest*, *Support Vector Machines (SVM)*, *Decision Tree*, *Neural Network*, dan *k-Nearest Neighbor (KNN)*.

2. *Unsupervised Learning*

Unsupervised Learning disebut juga *cluster* merupakan pembelajaran tak terarah atau tidak terawasi yang biasanya ditandai dengan kumpulan datanya tidak mempunyai atribut keputusan atau *class/label/target* pada himpunan data (Pradnyana & Agustini, 2022). Penggunaan metode ini bertujuan untuk mendalami dan mencari persamaan dari data yang dilatih serta mengidentifikasi struktur-struktur yang ada pada data tersebut (Nurhidayat, dkk., 2021). Algoritma *Unsupervised Learning* diantaranya *k-Mean*, *Aprior*, *Independent Subspace Analysis (ISA)*, dsb.

3. *Semi-Supervised Learning*

Semi-supervised learning merupakan penggabungan dari *supervised learning* dan *unsupervised learning* dengan mengelompokkan kumpulan data dengan atau tanpa *class/label/target* ke dalam beberapa *cluster*.

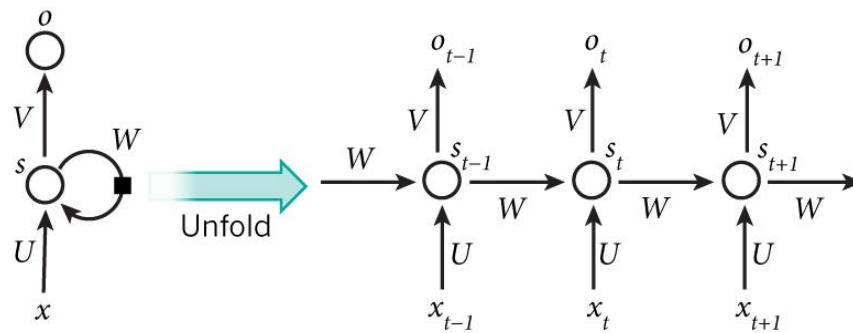
2.7 *Deep Learning*

Deep Learning merupakan bagian dari *Machine Learning* yang dibangun berdasarkan jaringan saraf tiruan (*Neural Network*) dalam mencari solusi penyelesaian suatu masalah tertentu. *Deep Learning* menggunakan beberapa lapisan diantara lapisan masukan (*input layer*) dan lapisan keluaran (*output layer*) yang dapat digunakan dalam melakukan pemrosesan nonlinear dengan beberapa tingkat dengan hasil keluarannya akan digunakan sebagai *feature learning* dan klasifikasi pola (*pattern classification*) (Diponegoro, dkk., 2021). Setiap lapisan menggunakan hasil keluaran dari lapisan sebelumnya sebagai masukannya. Kelebihan utama *Deep Learning* adalah mampu untuk merubah data nonlinear menjadi linear melalui serangkaian transformasi (*hidden layer*) (Putra, 2022). Implementasi algoritma *Deep Learning* dapat digunakan sebagai pengenalan pola (*pattern analysis*) dan klasifikasi, terutama dalam pengenalan gambar, suara, ucapan, teks dan deret waktu.

2.7.1 *Recurrent Neural Networks*

Recurrent Neural Networks merupakan salah satu arsitektur jaringan saraf tiruan dengan sistem kerja menyerupai kerja syaraf otak manusia dengan menggunakan fungsi iteratif (pengulangan) (Graves, 2012). *Recurrent Neural Networks* dirancang untuk memproses data berurutan atau berulang (*sequential data*) dan dapat digunakan dalam mengklasifikasikan, mengelompokkan dan membuat prediksi terutama terkait dengan data deret waktu. Jaringan RNN dapat menyimpan memori yang memungkinkan untuk mengenali pola data dengan baik dan membuat prediksi yang akurat dengan melakukan *looping* dalam arsitekturnya. *Recurrent Neural Networks* merupakan bagian dari *neural network* sehingga lapisan-lapisan yang terdapat di dalamnya adalah sebagai berikut (Runtu & Lina, 2022):

1. Masukan (*Input Layer*)
Merupakan lapisan penerima informasi yang akan diteruskan dari satu *neuron* ke *neuron* lain dalam jaringan.
2. Lapisan Tersembunyi (*Hidden Layer*)
Merupakan lapisan tersembunyi yang berfungsi untuk meningkatkan kemampuan jaringan dalam memecahkan permasalahan.
3. Keluaran (*Output Layer*)
Merupakan lapisan yang menghasilkan *output* berupa pemecahan masalah dari hasil pembelajaran.
4. Bobot (*Weight*)
Bobot merupakan salah satu parameter dalam jaringan saraf berulang.



Gambar 1. Struktur RNN
(Firmansyah dkk, 2020)

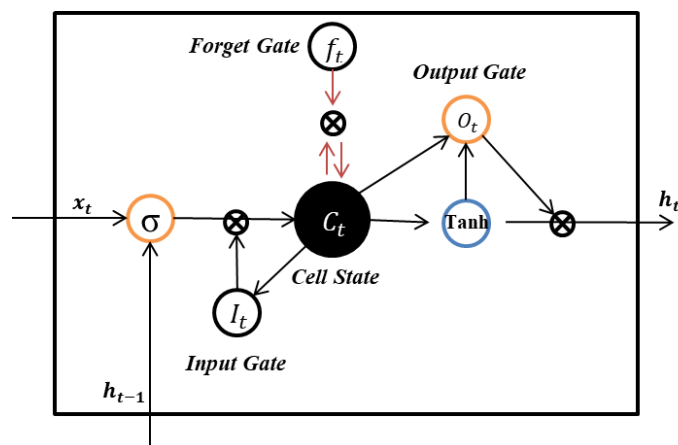
Berdasarkan Gambar 1, RNN akan memproses data secara sekuensial dan *hidden layer* akan menangkap informasi pada *time step* sebelumnya dimana *output* berasal dari waktu saat ini dan memori sebelumnya kemudian diteruskan menuju *hidden layer* berikutnya pada skala waktu selanjutnya. Sehingga dapat diartikan bahwa setiap *input* yang masuk dan menghasilkan *output*, *output* tersebut akan diproses kembali sebagai *input* baru untuk di proses di dalam *hidden layer* yang dilakukan secara berulang hingga memperoleh *output* yang sesuai target (Graves, 2012).

2.7.2 Long Short Term Memory

Long Short Term Memory merupakan salah satu pengembangan dari RNN yang dapat digunakan dalam pemodelan data *time series* dan memiliki kemampuan untuk mengolah informasi menjadi lebih akurat (Zhang, 2016). Metode LSTM cocok digunakan untuk data *time series* dalam mengklasifikasi, memproses dan membuat prediksi dalam menangkap suatu hubungan fungsional yang tidak diketahui (Yadav, dkk., 2020). Data yang masuk pada LSTM akan diproses secara berulang-ulang di setiap *output* dari *hidden layer* sehingga mendapatkan hasil yang paling akurat (Farhan, dkk., 2021). Tujuan utama dari algoritma LSTM dalam *forecasting* yaitu membuat prediksi terbaik yang didasarkan pada

semakin kecil tingkat kesalahan yang diperoleh maka akan semakin tepat suatu metode untuk memprediksikannya (Wiranda & Sadikin, 2019).

Metode LSTM dapat mengatasi *vanishing gradient* yang menjadi permasalahan dari RNN dalam memproses ketergantungan jangka panjang (*long term dependencies*) yang dapat mengurangi tingkat akurasi dari suatu prediksi (Wiranda & Sadikin, 2019). Selain itu, LSTM dapat mempertahankan *error* yang terjadi ketika melakukan *backpropagation* untuk menghasilkan ketepatan informasi lebih baik (Zhang, 2016). LSTM mempunyai *memory block* yang dapat menentukan nilai mana yang akan terpilih sebagai keluaran yang relevan terhadap masukan yang diberikan (Wiranda & Sadikin, 2019). Arsitektur LSTM secara umum terdiri dari *memory cell*, *input gate*, *output gate* dan *forget gate*.



Gambar 2. Arsitektur LSTM
(Chung & Shin, 2018)

Berikut ini merupakan struktur dari arsitektur LSTM (Runtu & Lina, 2022):

1. *Forget Gate*

Lapisan pertama dalam LSTM disebut sebagai *forget gate* yang bertujuan untuk memutuskan informasi mana yang kurang relevan dan tidak dibutuhkan oleh sebuah sistem, akan dilupakan atau dihapus menggunakan fungsi sigmoid. Data x_t merupakan *input* data (vektor *input* x dalam *timestep* t) dan

h_{t-1} merupakan vektor *hidden state* dalam *time step* sebelumnya $t - 1$. Perhitungan nilai *forget gate* adalah sebagai berikut:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_i) \quad (2.15)$$

dimana:

- f_t = *forget gate*
- σ = fungsi sigmoid
- W_t = nilai *weight* untuk *forget gate*
- h_{t-1} = nilai *output* dalam *timestep* $t - 1$
- x_t = nilai *output* dalam *timestep* t
- b_t = nilai bias *forget gate*

2. *Input Gate*

Lapisan kedua disebut *input gate* yang bertujuan untuk menseleksi dan menentukan informasi berguna ke bagian *cell state* menggunakan fungsi sigmoid. *Input gate* mempunyai 2 fungsi, fungsi pertama adalah menambahkan informasi sebelumnya yang telah diseleksi melewati *forget gate*. Berikut merupakan perhitungan nilai *input gate*:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.16)$$

dimana:

- i_t = *input gate*
- σ = fungsi sigmoid
- W_i = nilai *weight* untuk *input gate*
- h_{t-1} = nilai *output* dalam *timestep* $t - 1$
- x_i = nilai *input* dalam *timestep* t
- b_i = nilai bias *input gate*

Input gate akan membentuk kandidat vektor baru \tilde{C}_t menggunakan fungsi aktivasi tanh agar dapat mengontrol berapa banyak informasi baru ditambahkan. Perhitungan yang dilakukan adalah sebagai berikut:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.17)$$

dimana:

\tilde{C}_t = nilai baru yang ditambahkan pada *cell state*

\tanh = fungsi tanh

W_c = nilai *weight* untuk *cell state*

h_{t-1} = nilai *output* dalam *timestep* $t - 1$

x_t = nilai *input* dalam *timestep* t

b_c = nilai bias *cell state*

Selanjutnya yaitu memperbarui nilai *cell state* lama C_{t-1} menjadi *cell state* baru C_t menggunakan fungsi berikut:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.18)$$

dimana:

C_t = nilai *cell state* baru

C_{t-1} = *cell state* sebelumnya $t - 1$

\tilde{C}_t = nilai baru yang ditambahkan pada *cell state*

f_t = *forget gate*

i_t = *input gate*

3. Output Gate

Lapisan terakhir LSTM adalah *output gate* yang bertujuan untuk menjalankan aktivasi sigmoid untuk menghasilkan nilai *output* pada *hidden state* dan menempatkan *cell state* pada tanh. Hasil aktivasi nilai *output* sigmoid dan *output* tanh selanjutnya akan dilakukan perkalian sebelum ke tahap berikutnya. Berikut merupakan perhitungannya:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.19)$$

dimana:

o_t = *output gate*

σ = fungsi sigmoid

W_o = nilai *weight* untuk *output gate*

h_{t-1} = nilai *output* dalam *timestep* $t - 1$
 x_i = nilai *input* dalam *timestep* t
 b_o = nilai bias *output gate*

Nilai *output* terakhir dapat didefinisikan sebagai berikut:

$$h_t = o_c * \tanh(C_t) \quad (2.20)$$

dimana:

h_t = nilai *output* ke- t
 o_t = *output gate*
 \tanh = fungsi tanh
 C_t = *cell state*

2.7.3 Fungsi Aktivasi

Fungsi aktivasi dalam *neural network* digunakan untuk memutuskan setiap *neuron* dalam jaringan diaktifkan atau tidak dan dapat digunakan dalam menormalisasikan *output* setiap *neuron* mempunyai nilai dalam *range* tertentu. Adapun fungsi aktivasi yang digunakan dalam LSTM, yaitu:

1. Fungsi sigmoid (fungsi logistic)

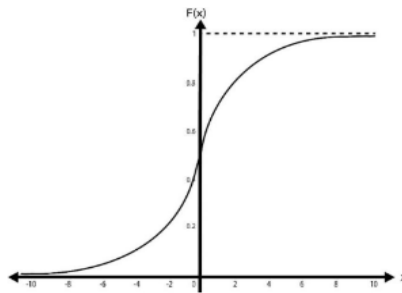
Fungsi aktivasi sigmoid digunakan untuk menormalisasikan *output* dengan mentransformasikan nilai yang berada dalam rentang -1 dan 1 menjadi nilai diantara 0 dan 1 (Putra, dkk., 2022). Berikut merupakan persamaan dari fungsi sigmoid:

$$\sigma(x) = \frac{1}{(1 + e^{-x})} \quad (2.21)$$

dimana:

e = bilangan euler
 x = data

Adapun grafik yang dihasilkan dari fungsi sigmoid sebagai berikut:



Gambar 3. Grafik Fungsi Sigmoid (Logistik)
(Akbar dkk, 2022)

2. Fungsi tan hiperbolik (tanh)

Fungsi aktivasi tanh digunakan untuk menormalisasikan *output* pada setiap *neuron* yang dilalui memiliki nilai diantara -1 dan 1 (Putra, dkk., 2022).

Berikut merupakan persamaan dari fungsi tanh:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.22)$$

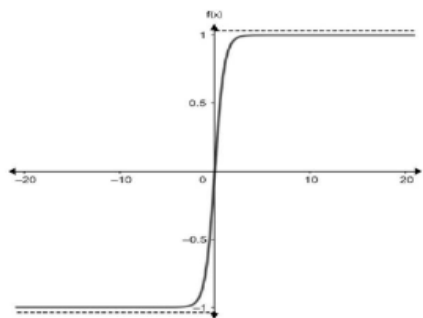
$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} \quad (2.23)$$

dimana:

e = bilangan euler

x = data

Adapun grafik yang dihasilkan dari fungsi tanh sebagai berikut:



Gambar 4. Grafik Fungsi Tanh
(Akbar dkk, 2022)

2.7.4 Hyperparameter

Pemilihan parameter yang tepat pada model dilakukan untuk meningkatkan hasil yang maksimal. *Hyperparameter Tuning* merupakan parameter yang berperan dalam pembelajaran mesin (*machine learning*) dan algoritma *deep learning* karena parameter yang dihasilkan akan mempengaruhi kinerja model agar didapatkan model dengan performa yang optimal (Putra, dkk., 2022). Adapun parameter yang digunakan yaitu:

1. *Epoch*

Epoch merupakan parameter yang digunakan dalam menentukan berapa kali *neural network* melakukan proses pelatihan untuk mengolah keseluruhan *dataset*. Satu *epoch* artinya ketika seluruh *dataset* sudah melalui proses pelatihan pada *neural network* sampai dikembalikan lagi ke awal. Proses pelatihan model tidak dapat dilakukan hanya menggunakan satu *epoch*, dikarenakan *dataset* yang digunakan terbatas dan untuk mengoptimalkan grafik *gradient descent* memerlukan proses iteratif. (Putra, dkk., 2022). Proses pelatihan satu *epoch* dapat berlangsung cukup lama, maka diperlukan pembagian data ke dalam setiap batch yang dikenal dengan istilah *batch size*.

2. *Batch Size*

Batch size merupakan jumlah berapa banyak sampel dalam *dataset* yang dimasukkan ke dalam *neural network* sebelum bobot disesuaikan akan digunakan dalam satu iterasi (Putra, dkk., 2022).

3. *Adaptive Moment Estimation (ADAM)*

Adaptive Moment Estimation (ADAM) merupakan algoritma optimasi dari penggabungan *Adaptive Gradient Algorithm (AdaGrad)* dan *Root Mean Square Propagation (RMSProp)*. *Adaptive Moment Estimation (ADAM)* mampu untuk memberikan pengoptimalan suatu algoritma yang mampu menangani *sparse gradients* pada *noisy problem* (Witanto, dkk., 2022). Sehingga metode ini efisien digunakan ketika bekerja pada data dan parameter yang besar .

2.8 *Scaling Data*

Normalisasi bertujuan untuk menghilangkan data yang berangkap, mempercepat model, dan meminimalkan *error* dalam proses pembelajaran dengan membagi skala nilai data tersebut ke dalam suatu *range* interval (Lattifia, dkk. 2022). Metode normalisasi data yang dapat digunakan, yaitu:

1. *Min-Max Scaler*

Min-Max Scaler merupakan metode yang dilakukan untuk normalisasi dengan melakukan transformasi linear dengan membagi skala nilai data ke dalam interval 0 hingga 1 sehingga menghasilkan keseimbangan nilai perbandingan antar *dataset*. Persamaan untuk *Min-Max Scaler* dapat ditulis sebagai berikut (Larose, 2014):

$$X_n = \frac{X_0 - X_{min}}{X_{max} - X_{min}} \quad (2.24)$$

dimana:

- X_0 = nilai data aktual
- X_n = nilai hasil normalisasi
- X_{min} = nilai minimum dari data aktual
- X_{max} = nilai maksimum dari data aktual

2. *Standar Scaler*

Standard Scaler merupakan metode yang dilakukan untuk normalisasi data dengan mentransformasi ke dalam bentuk distribusi dengan nilai *mean* (μ) yaitu 0 dan nilai standar deviasi (σ) yaitu 1. Persamaan untuk *Standar Scaler* dapat ditulis sebagai berikut (Larose, 2014):

$$Z = \frac{X_i - \bar{X}}{\sigma} \quad (2.25)$$

dimana:

- Z = nilai hasil normalisasi

X_i	= nilai data aktual
\bar{X}	= nilai minimum
σ	= nilai standar deviasi

2.9 Evaluasi Model

Evaluasi model digunakan untuk memverifikasi keakuratan prediksinya dengan membandingkan nilai prediksi dengan nilai pengamatan atau nilai sebenarnya (Chen, dkk., 2018). Prediksi terbaik dilihat berdasarkan tingkat keakuratan prediksinya, semakin kecil tingkat kesalahan yang dihasilkan menunjukkan bahwa semakin akurat sebuah metode dalam prediksi. Beberapa evaluasi model yang dapat digunakan, diantaranya:

1. *Mean Absolute Percent Error* (MAPE)

Mean Absolute Percent Error merupakan persentase *error* yang dihasilkan dari nilai minimum selisih antara data aktual dengan data hasil prediksi yang dinyatakan sebagai persentase nilai aktual (Wei, 2006).

$$MAPE = \sum_{i=1}^n \frac{|A_i - F_i|}{A_i} \times 100\% \quad (2.26)$$

dimana:

A_i	= nilai data aktual atau pengamatan
F_i	= nilai hasil prediksi
n	= banyaknya data
i	= indeks waktu, $i = 1, 2, \dots$

Penggunaan MAPE pada evaluasi dapat melihat tingkat ketepatan akurasi terhadap besarnya nilai aktual dengan hasil prediksinya. Adapun kriteria MAPE ditunjukkan pada Tabel 1 (Lewis, 1982):

Tabel 1. Kriteria nilai MAPE

Nilai MAPE	Model Prediksi
<10%	Sangat baik
10% - 20%	Baik
20% - 50%	Cukup baik
>50%	Buruk

2. Mean Squared Error (MSE)

Mean Squared Error merupakan minimum dari kesalahan prediksi yang di kuadratkan. MSE merupakan metode lain yang digunakan untuk mengevaluasi kesalahan pada prediksinya (Wei, 2006). Nilai MSE dikatakan semakin akurat hasil prediksinya apabila semakin kecil atau mendekati nol.

$$MSE = \frac{\sum_{i=1}^n (A_i - F_i)^2}{n} \quad (2.27)$$

dimana:

A_i = nilai data aktual atau pengamatan

F_i = nilai hasil prediksi

n = banyaknya data

i = indeks waktu, $i = 1, 2, \dots$

3. Root Mean Square Error (RMSE)

Root Mean Square Error merupakan akar dari minimum dari kesalahan prediksi yang di kuadratkan. RMSE merupakan besarnya tingkat kesalahan hasil prediksi. Nilai RMSE yang mendekati nol mengindikasikan hasil prediksi semakin baik (Lattifia, 2022).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - F_i)^2}{n}} \quad (2.28)$$

dimana:

A_i = nilai data aktual atau pengamatan

F_i = nilai hasil prediksi

n = banyaknya data

i = indeks waktu, $i = 1, 2, \dots$

2.10 Hybrid SARIMA-LSTM

Metode *hybrid* merupakan perkembangan suatu metode dua atau lebih model tunggal dalam suatu sistem. Model *hybrid* SARIMA–LSTM merupakan model *hybrid* yang dapat menyelesaikan masalah linear dan nonlinear (Tahyudin, dkk., 2022). Model SARIMA dan LSTM memiliki kelebihan dan kekurangan masing-masing. Model SARIMA baik digunakan untuk memproses data *time series* yang bersifat linear dan memiliki komponen musiman serta memiliki akurasi yang tinggi saat digunakan untuk peramalan jangka pendek. Akan tetapi, model SARIMA akan kesulitan dalam memproses data yang bersifat nonlinear dan jika digunakan untuk peramalan jangka panjang, akurasi peramalannya akan menurun. Sedangkan model LSTM mampu dalam memproses data bersifat linear maupun nonlinear, namun dalam memproses data akan membutuhkan waktu yang lama. Model LSTM memiliki *memory cell* yang dapat menyimpan informasi dalam jangka waktu panjang sehingga baik digunakan untuk peramalan jangka panjang.

Sebuah model gabungan (*hybrid*) dirancang untuk saling menutupi kekurangan dari tiap model dan memanfaatkan kelebihan masing-masing model sehingga dapat meningkatkan akurasi dari hasil peramalannya (Rowan, dkk., 2022). Kombinasi dari model *time series* yang memiliki struktur autokorelasi linear dan nonlinear secara umum dapat ditulis sebagai (Zhang, 2003):

$$Z_t = L_t + N_t \quad (2.29)$$

dimana:

- Z_t = data aktual ke- t
- L_t = komponen linear ke- t
- N_t = komponen nonlinear ke- t
- t = indeks waktu

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil Tahun Akademik 2022/2023, bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung

3.2 Spesifikasi Perangkat Penelitian

Spesifikasi perangkat yang digunakan dalam penelitian :

Jenis : Laptop
Merk/Tipe : Acer Aspire E14
Prosesor : Intel Celeron (2.16 Hz, Dual Core) N2830
Tipe Sistem : Windows 10 Home
Memory : 4 GB DDR4 SDRAM
Software : *Python 3.0*

3.3 Data Penelitian

Data yang digunakan merupakan data cuaca harian yang tersedia pada situs BMKG <http://dataonline.bmkg.go.id/home>. Data yang tersedia memiliki 9 variabel dengan data yang digunakan pada penelitian ini adalah data historis dari

suhu minimum (T_n) cuaca harian yang diambil bertempat di Stasiun Meteorologi Maritim Panjang, Tanjung Karang, Kota Bandar Lampung selama 13 tahun dan terdapat 4.746 data. Data suhu minimum di Kota Bandar Lampung yang digunakan berkaitan dengan metode SARIMA yang merupakan metode *univariat* sehingga hanya satu variabel yang digunakan dan memiliki pola musiman.

Tabel 2. Data Penelitian

Tanggal	T_n	T_x	T_{avg}	RH_{avg}	RR	ss	ff_x	ddd_x	ff_{avg}
4/1/2010	21.0	34.0	29.6	70.0	Nan	5.0	4.0	270.0	0.0
5/1/2010	21.0	32.0	28.9	69.0	NaN	0.4	6.0	270.0	0.0
6/1/2010	21.0	32.0	28.7	73.0	NaN	0.0	5.0	360.0	0.0
...
31/12/2022	24,8	31,4	26,5	83	0	0,8	4	180	1
1/1/2023	24,2	30,2	27,5	82	56,8	3,6	4	360	1

dimana:

T_n = suhu minimum ($^{\circ}C$)

T_x = suhu maksimum ($^{\circ}C$)

T_{avg} = suhu rata-rata ($^{\circ}C$)

RG_{avg} = kelembaban minimum (%)

RR = curah hujan (mm)

ss = durasi penyinaran matahari (jam)

ff_x = kecepatan angin maksimum (m/s)

ddd_x = arah angin dengan kecepatan maksimum ($^{\circ}$)

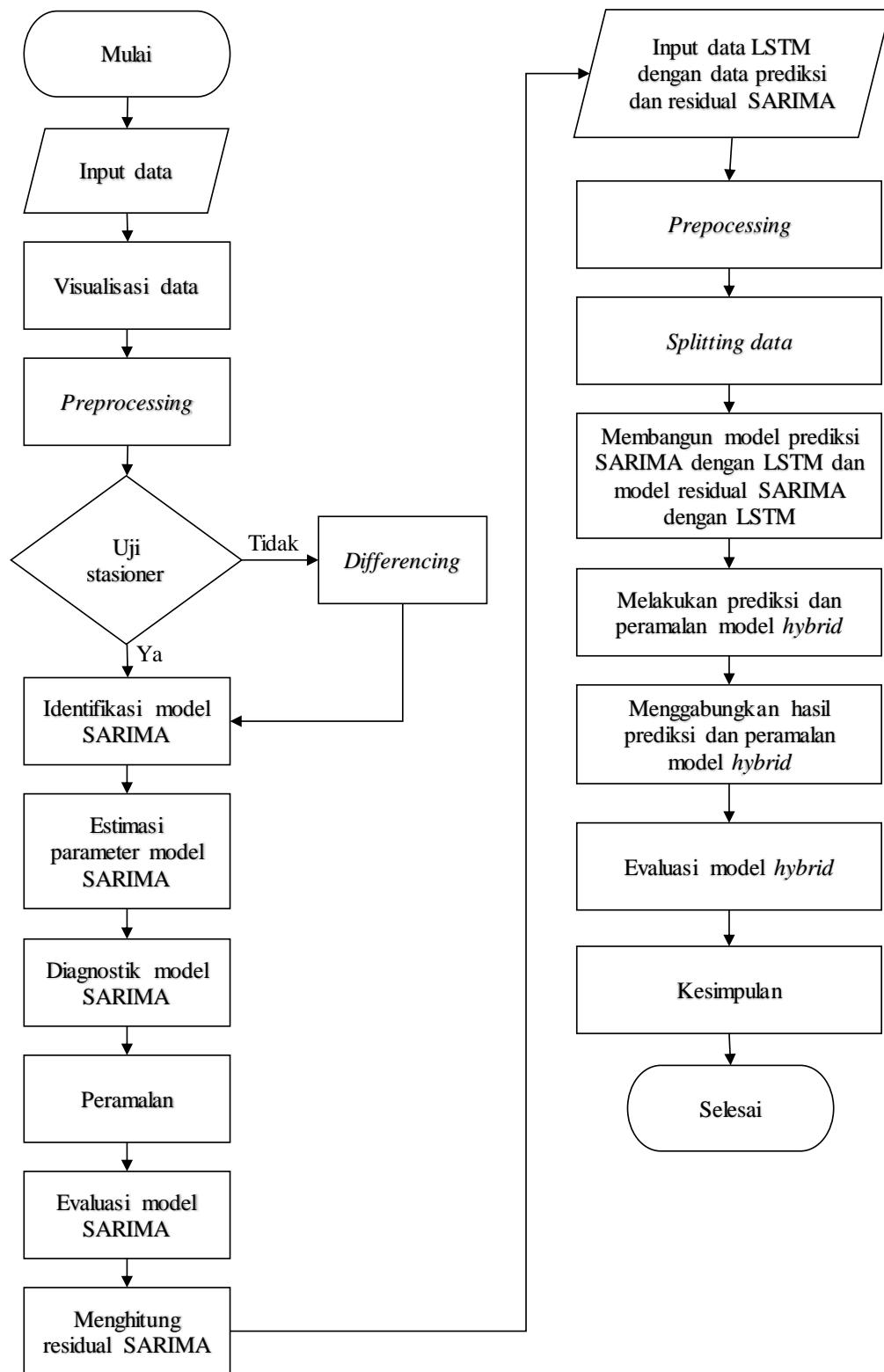
ff_{avg} = kecepatan angin minimum ($^{\circ}$)

3.4 Metode Penelitian

Berikut ini merupakan alur dari pengerjaan metode *hybrid* SARIMA – LSTM:

1. Melakukan studi literatur mengenai proses model *hybrid* SARIMA – LSTM dari berbagai sumber, seperti jurnal, buku dan dibantu oleh narasumber yang memahami metode *hybrid* SARIMA – LSTM.
2. Mengumpulkan data yang akan digunakan dalam proses peralaman dengan metode *hybrid* SARIMA – LSTM.
3. Melakukan *input* data cuaca di Kota Bandar Lampung yang didapat dari *website* BMKG.
4. Melakukan *preprecoessing* data dengan mengecek *missing value* dan melakukan konversi data *input* harian menjadi mingguan.
5. Melakukan visualisasi data dengan *plotting* untuk melihat pola dekomposisi dalam data.
6. Melakukan uji stasioneritas untuk data non-musiman dengan menggunakan uji *Augmented Dickey-Fuller* (ADF). Jika data belum memenuhi uji stasioneritas maka akan dilakukan proses *differencing*.
7. Melakukan uji stasioneritas untuk data musiman dengan menggunakan uji *Augmented Dickey-Fuller* (ADF).
8. Melakukan identifikasi untuk model ARIMA non-musiman dengan orde p, d, q dan musiman dengan orde P, D, Q dan S berdasarkan plot ACF dan PACF.
9. Menentukan estimasi model SARIMA sementara.
10. Melakukan uji diagnostik model untuk mengetahui apakah model tersebut sudah cukup baik untuk digunakan dalam prediksi dengan memenuhi asumsi *white noise*.
11. Melakukan prediksi dan peramalan model SARIMA menggunakan model terbaik SARIMA serta menghitung nilai residual dengan cara mengurangi data aktual dengan data hasil peramalan model SARIMA.
12. Melakukan evaluasi model prediksi SARIMA.
13. Data prediksi dan data residual dari model SARIMA digunakan sebagai input yang akan diolah dengan menggunakan metode LSTM.

14. Melakukan *preprocessing* data, yaitu permasalahan *missing value* dan konversi data harian menjadi data mingguan.
15. Melakukan normalisasi model prediksi dan model residual menggunakan *MinMaxScaler*.
16. Melakukan *splitting* untuk kedua model dengan skema 70% *training* 30% *testing* dan 80% *training* 20% *testing*.
17. Membangun dua model LSTM dengan menggunakan metode LSTM sebagai landasan dalam melakukan proses *hybrid*. Model pertama yaitu model prediksi SARIMA dengan LSTM merupakan model yang akan digunakan untuk memprediksi data prediksi SARIMA. Sedangkan model kedua, model residual SARIMA dengan LSTM merupakan model yang digunakan untuk memprediksi data residual dari SARIMA.
18. Penentuan parameter model LSTM untuk data prediksi dan data residual dari metode SARIMA yang dilakukan dengan menggunakan *hypertuning* untuk menentukan parameter terbaik.
19. Melakukan prediksi dan peramalan dengan model prediksi SARIMA dengan LSTM
20. Melakukan prediksi dan peramalan dengan model residual SARIMA dengan LSTM
21. Menggabungkan hasil prediksi dan peramalan dari kedua model *hybrid* SARIMA-LSTM dengan proses penjumlahan.
22. Melakukan evaluasi model prediksi *hybrid* SARIMA-LSTM.



Gambar 5. Diagram Alir Metode Penelitian

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Metode *hybrid* SARIMA-LSTM merupakan metode yang dapat digunakan dalam melakukan peramalan terkait cuaca dengan baik. Adapun kesimpulan yang dapat diambil selama proses penelitian terhadap cuaca di Kota Bandar Lampung dengan metode *hybrid* SARIMA-LSTM adalah sebagai berikut:

1. Metode SARIMA melakukan prediksi menggunakan model SARIMA terbaik, yaitu model $ARIMA(6,1,0)(3,1,0,26)$ yang menghasilkan *output* prediksi dan residual dari model SARIMA. Hasil tersebut akan di *hybrid*-dengan metode LSTM dengan parameter *LSTM Unit* dan *Batch Size* yang di *hypertunning* untuk mendapatkan model terbaik.
2. Metode ini menggunakan 2 skema *splitting data* yang berbeda, dimana pada penggunaan *splitting data* 80% *training* dan 20% *testing* menghasilkan nilai performa yang lebih baik dibandingkan dengan skema *splitting data* 70% *training* dan 30% *testing*.
3. Metode *hybrid* SARIMA-LSTM merupakan metode yang baik digunakan dalam melakukan prediksi untuk data yang memiliki periode musiman seperti data suhu minimum. Pada metode ini menghasilkan nilai akurasi MSE sebesar 0.1174, RMSE sebesar 0.3426 dan MAPE sebesar 0,0104 %.
4. Hasil dari prediksi akan diproses untuk mencari peramalan untuk 21 minggu kedepan. Peramalan menggunakan metode *hybrid* SARIMA-LSTM menghasilkan peramalan yang lebih baik dan mengikuti data terkini yang tersedia. Sehingga metode *hybrid* SARIMA-LSTM baik digunakan untuk melakukan peramalan.

5.2 Saran

Pada penelitian selanjutnya dapat menggunakan variabel lain untuk melakukan prediksi dan peramalan. Penelitian selanjutnya juga dapat mencoba untuk menggunakan lebih dari 1 variabel *eksogenous* yang dapat di analisis, seperti dengan menggunakan metode SARIMAX yang di *hybrid* dengan LSTM.

DAFTAR PUSTAKA

- Aksan, I. & Khalilah, N. 2020. Aplikasi Metode Arima Box-Jenkins untuk Meramalkan Penggunaan Harian Data Seluler. *Journal of Mathematics: Theory and Applications*. **2**(1): 5-10.
- Akbar, K., Hayaty, M. 2020. Data Balancing untuk Mengatasi Imbalance Dataset pada Prediksi Produksi Padi. *Jurnal Ilmiah Intech Information Technology Journal of UMUS*. **2**(2): 1-14.
- Badan Meteorologi Klimatologi dan Geofisika (BMKG). 2021. *Buku Saku Klimatologi*. BMKG, Jakarta.
- Badan Perencanaan Pembangunan Nasional (Bappenas). 2010. *Indonesia Climate Change Sektoral Roadmap (ICCSR) Sektor Kesehatan*. Bappenas, 2010.
- Chen, P., Niu, A., Liu, D., Jiang, W. & Ma, B. 2018. Time Series Forecasting of temperatures using SARIMA: An Example from Nanjing. *IOP Conference Series: Materials Science and Engineering*. **394**(5): 1-7.
- Chung, H., Shin, K. 2018. Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction. *MDPI*. **10**(10):1-18
- Ding, X. W. 2022. A Time Series-Based Statistical Approach for Trade Turnover Forecasting and Assessing: Evidence from China and Russia. *Journal of Asian Finance and Business*. **9**(4): 83-92.
- Diponegoro, M. H., Kusumawardani, S. S. & Hidayah, I. 2021. Tinjauan Pustaka Sistematis: Implementasi Metode Deep Learning pada Prediksi Kinerja Murid. *Jurnal Nasional Teknik Elektro dan Teknologi Terapan*. **10**(2): 131-138.
- Farhan, A., Prasasti, A. L. & Parywasto, W. 2021. Implementasi Recurrent Neural Network dalam Memprediksi Kepadatan Restoran Berbasis LSTM. *Jurnal Media Informatika Budiman*. **5**(2): 74-80. neneural networks

- Firmansyah, M. R., Ilyas, R., Kasyidi, F. 2020. Klasifikasi Kalimat Ilmiah Menggunakan Recurrent Neural Network. Hlm. 488-495. *Prosiding The 11th Industrial Research Workshop and National Seminar*. Bandung
- Graves, A. 2012. *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer, Berlin.
- Hand, J., Kamber, M. & Pei, J. 2012. *Data Mining Concepts and Techniques. Third Edition*. Morgan Kaufmann Publishers, USA.
- Hand, D. H., Mannila, H. & Smyth, P. 2001. *Principles of Data Mining*. MIT Press, Cambridge.
- Larose, D. T. & Larose, C. D. 2014. *Discovering Knowledge in Data – An Introduction to Data Mining*. Second Edition. John Wiley and Sons, New Jersey.
- Lewis, C. D. 1982. *Industrial and Business Forecasting Methods*. Butterworth, London.
- Lattifia, T., Buana, P. W. & Rusjyanthi, N. K. D. 2022. Model Prediksi Cuaca Menggunakan Metode LSTM. *Jurnal Ilmiah Teknologi dan Komputer*. 3(1): 994-1000.
- Makridakis, S., Wheelright, S.C., and McGee V. E. 1995. *Forecasting: Methods and Applications*. Second Edition. John Wiley and Sons, New Jersey.
- Montgomery, D. C., Jennings, C. L. & Kulahci, M. 2015. *Introduction to Time Series Analysis and Forecasting*. Second Edition. John Wiley and Sons, New Jersey.
- Muntiari, N. R. & Hanif, K. H. 2022. Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning. *Jurnal Ilmu Komputer dan Teknologi*. 3(1): 1-6.
- Nurhidayat, A. I., Asmunin & Fatrianto, D. 2021. Prediksi Akademik Mahasiswa Menggunakan Machine Learning dengan Sequential Minimal Optimization untuk Pengelola Program Studi. *Journal Information Engineering and Education Technology*. 5(2): 84-91.
- Padhilha, G. A. G., Ko, J., Jung, J. J. & Neto, P. S. G. M. 2022. Transformer-Based Hybrid Forecasting Model for Multivariate Renewable Energy. *Applied Sciences*. 12(21): 1-13.
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z. & Barrow, D. K.. 2022. Forecasting: theory and practice. *International Journal of Forecasting*. 38(3): 705-871.

- Pradnyana, G. A & Agustini, K. 2022. *Konsep Dasar Data Mining*. Edisi Pertama. Universitas Terbuka, Tangerang Selatan.
- Putra, T. I. Z. M. & Suprpto & Bukhori, A. F. 2022. Model Klasifikasi Berbasis Multiclass Classification dengan Kombinasi Indobert Embedding dan Long Short-Term Memory untuk Tweet Berbahasa Indonesia. *Jurnal Ilmu Siber dan Teknologi Digital*. **1**(1): 1-28.
- Rowan, Muflikhah, L. & Cholissodin I. 2022. Peramalan Kasus Positif COVID-19 di Jawa Timur menggunakan metode Hybrid ARIMA-LSTM. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. **6**(9): 4146-4153.
- Runtu, K., D. & Lina. 2022. Pengenalan Aktivitas Manusia Di Supermarket Dengan Metode Long Short Term Memory. *Jurnal Ilmu Komputer dan Sistem Informatika*. **10**(2): 1-5.
- Sirisha, U. M., Belavagi, M. C. & Attigeri, A. G. 2022. Profit Predictions Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison. *IEEE Access*. **10**: 124715-124727.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning*. **15**(56): 1929-1958.
- Sun, Q., Wan, J. & Liu, S. Estimation of Sea Level Variability in the China Sea and Its Vicinity Using SARIMA and LSTM Models. *IEEE Journal*. **99**(13): 3317-2020.
- Tahyudin, I., Wahyudi, R. & Nambo, H. 2021. SARIMA-LSTM Combination for Covid-19 Case Modeling. *IJUM Engineering Journal*. **23**(2): 171-182.
- Wei, W. W. S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition. Pearson Education, America.
- Wiranda, L. & Sadikin, M. 2019. Penerapan Long Short Term Memory pada Data Time Series untuk Memprediksi Penjualan Produk PT. Mediska Farma. *Jurnal Nasional Pendidikan Teknik Informatika*. **8**(2): 184-196.
- Witanto, K. S., Sanjaya, N. A., Karyawati, E., Kadyanan., I. G. A. G. A., Suhartana, I. K. G. S. & Astuti, L. G. 2022. Implementasi LSTM pada Analisis Sentimen Review Film Menggunakan Adam dan RMSprop Optimizer. *Jurnal Elektronik Ilmu Komputer Udayana*. **10**(4): 351-362.
- Wu, D. C. W., He, L. J. K. & Tso, K. F. G. 2021. Forecasting Tourist Daily Arrivals With a Hybrid SARIMA-LSTM Approach. *Journal of Hospitality & Tourism Research*. **45**(1): 52-67.

- Yadav, A., Jha, C. K. & Sharan, A. 2020. Optimizing LSTM for time series prediction in Indian stock market. *Procedia Computer Science*. **167**(4): 2091-2100.
- Zhang, G. P. 2003. *Neural Networks in Business Forecasting*. Idea Group Inc., USA.
- Zhang, L., Song, X. & Wu Y. 2016. *Theory, Methodology, Tools, and Applications for Modeling and Simulation of Complex Systems*. Springer, Singapore.