

**IMPLEMENTASI ALGORITMA BI-LSTM UNTUK PREDIKSI
CUACA BERBASIS DATA HISTORIS**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

FAUZ SYABANA PUTRA

21.11.4288

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**IMPLEMENTASI ALGORITMA BI-LSTM UNTUK PREDIKSI
CUACA BERBASIS DATA HISTORIS**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

FAUZ SYABANA PUTRA

21.11.4288

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA BI-LSTM UNTUK PREDIKSI CUACA
BERBASIS DATA HISTORIS**

yang disusun dan diajukan oleh

Fauz Syabana Putra

21.11.4288

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Februari 2025

Dosen Pembimbing,


Majid Rahardi, S.Kom., M.Eng
NIK. 190302393

HALAMAN PENGESAHAN
SKRIPSI
IMPLEMENTASI ALGORITMA BI-LSTM UNTUK PREDIKSI CUACA
BERBASIS DATA HISTORIS

yang disusun dan diajukan oleh

Fauz Syabana Putra

21.11.4288

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dr. Ferry Wahyu Wibowo, S.Si., M.Cs.
NIK. 190302235

Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fauz Syabana Putra
NIM : 21.11.4288

Menyatakan bahwa Skripsi dengan judul berikut:

Implementasi Algoritma Bi-LSTM untuk Prediksi Cuaca Berbasis Data Historis

Dosen Pembimbing : Majid Rahardi, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Februari 2025

Yang Menyatakan,



Fauz Syabana Putra

HALAMAN PERSEMBAHAN

Skripsi ini dipersembahkan dengan penuh rasa syukur kepada:

1. Allah SWT, atas segala rahmat, hidayah, dan kemudahan yang diberikan dalam menyelesaikan penelitian ini.
2. Orang tua tercinta, yang selalu memberikan doa, dukungan, dan kasih sayang tanpa batas.
3. Keluarga dan sahabat, yang senantiasa memberikan semangat dan motivasi selama proses penelitian ini.
4. Dosen pembimbing dan penguji, atas bimbingan, ilmu, serta arahnya dalam menyelesaikan skripsi ini.
5. Almamater, sebagai tempat menimba ilmu dan berkembang dalam bidang akademik serta profesional.

Semoga karya ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan menjadi langkah awal untuk kontribusi lebih lanjut di bidang penelitian.

KATA PENGANTAR

Puji dan syukur senantiasa dipanjatkan ke hadirat Allah SWT atas segala rahmat dan hidayah-Nya sehingga skripsi ini dapat diselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Informatika, Fakultas Komputer, Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, penulis mendapatkan banyak dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis menyampaikan apresiasi yang sebesar-besarnya kepada:

1. Allah SWT karena dengan rahmatnya dan hidayahnya sehingga penulis dapat menyelesaikan penelitian Tugas Akhir ini.
2. Kedua orang tua tercinta, yang selalu memberikan doa, dukungan, serta motivasi yang tiada henti.
3. Bapak Majid Rahardi, S.Kom., M.Eng selaku Dosen Pembimbing, yang dengan penuh kesabaran telah memberikan bimbingan, arahan, dan masukan berharga dalam penyelesaian skripsi ini.
4. Bapak Dr. Ferry Wahyu Wibowo, S.Si., M.Cs., Bapak Mulia Sulistiyono, S.Kom., M.Kom. Bapak Majid Rahardi, S.Kom., M.Eng selaku Tim Penguji, yang telah memberikan saran dan kritik yang membangun demi kesempurnaan penelitian ini.
5. Seluruh keluarga, sahabat, dan rekan mahasiswa, yang telah memberikan semangat, bantuan, dan kebersamaan selama proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan untuk perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan serta menjadi referensi bagi penelitian selanjutnya.

Yogyakarta, 18 Februari 2025

Fauz Syabana Putra

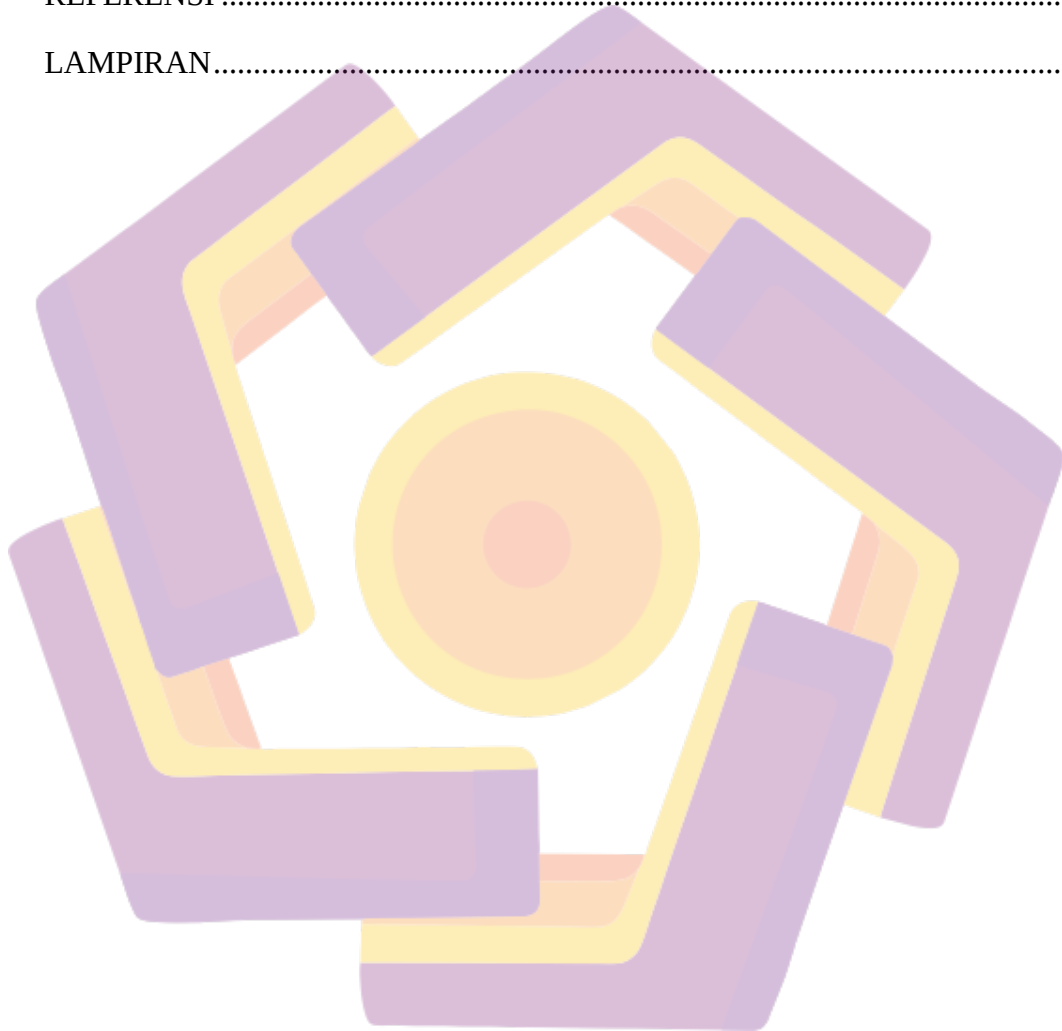
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xv
INTISARI	xvii
ABSTRACT.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	14

2.2.1 Prediksi Cuaca Berbasis Data Historis	14
2.2.2 Deep Learning.....	14
2.2.3 Bidirectional Long Short-Term Memory (Bi-LSTM)	14
2.2.4 Parameter Meteorologi.....	15
2.2.5 Preprocessing Data.....	16
2.2.6 TensorFlow dan Keras	16
2.2.7 Pembagian Data (Data Splitting)	17
2.2.8 Epoch	18
2.2.9 Optimasi Model	19
2.2.10 Evaluasi Model	19
BAB III METODE PENELITIAN	21
3.1 Objek Penelitian	21
3.2 Alur Penelitian.....	22
3.2.1 Akuisi Data	22
3.2.2 Import Library.....	23
3.2.3 Setel SEED.....	24
3.2.4 Memahami Variabel Data	24
3.2.5 Memeriksa Nilai Yang Hilang	25
3.2.6 Filter Negara Asia Tenggara.....	26
3.2.7 Encode Kondisi Cuaca	26
3.2.8 Memilih Fitur	27
3.2.9 Scaling Data	28
3.2.10 Membagi Data Menjadi Set Pelatihan Dan Pengujian.....	28
3.2.11 Membangun Model Bi-LSTM	29
3.2.12 Kompilasi Model (Optimizer : Adam).....	29

3.2.13	Melatih Model Dengan Early Stopping	30
3.2.14	Melakukan Hyperparameter Tuning	31
3.2.15	Evaluasi Model	32
3.2.16	Visualisasi Hasil.....	32
3.2.17	Membuat Simulasi Prediksi	32
3.2.18	Menampilkan Prediksi Kondisi Cuaca	32
3.3	Alat dan Bahan	33
3.3.1	Data Penelitian	33
3.3.2	Alat/Instrumen	36
BAB IV HASIL DAN PEMBAHASAN		39
4.1	Preprocessing	39
4.1.1	Memfilter Negara.....	39
4.1.2	Pemilihan Fitur.....	39
4.1.3	Encoding Kondisi Cuaca.....	43
4.1.4	Normalisasi Fitur dengan MinMaxScaler.....	44
4.1.5	Memecah Data Menjadi Data Training dan Data Testing	45
4.1.6	Reshape Data Training dan Testing.....	46
4.2	Hasil Pelatihan dan Evaluasi Model	47
4.2.1	Metrik Evaluasi.....	47
4.2.2	Grafik Loss.....	48
4.2.3	Grafik Accuracy	49
4.2.4	Confusion Matrix	50
4.2.5	Classification Report.....	52
4.3	Hyperparameter Tuning	54
4.3.1	Grafik Actual vs Predicted.....	56

4.4 Simulasi Prediksi	56
BAB V PENUTUP	59
5.1 Kesimpulan.....	59
5.2 Saran.....	60
REFERENSI	62
LAMPIRAN.....	67



DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	9
Tabel 4.1. Informasi Dataset Sebelum dan Setelah Preprocessing	37
Tabel 4.2. Fitur dalam Dataset Sebelum Pemilihan Fitur	37
Tabel 4.3. Fitur dalam Dataset Setelah Pemilihan Fitur	39
Tabel 4.4. Hasil Encoding Label Kondisi Cuaca	41
Tabel 4.5. Data Sebelum dan Sesudah Normalisasi	42
Tabel 4.6. Pembagian Data Training dan Testing	43
Tabel 4.7. Tabel Hasil Reshape dengan Data Training dan Testing	44
Tabel 4.8. Hasil Evaluasi Model	45
Tabel 4.9. Classification Report	51
Tabel 4.10. Sebelum dan Sesudah Hyperparameter Tuning	54
Tabel 4.11. Input Parameter Cuaca	57

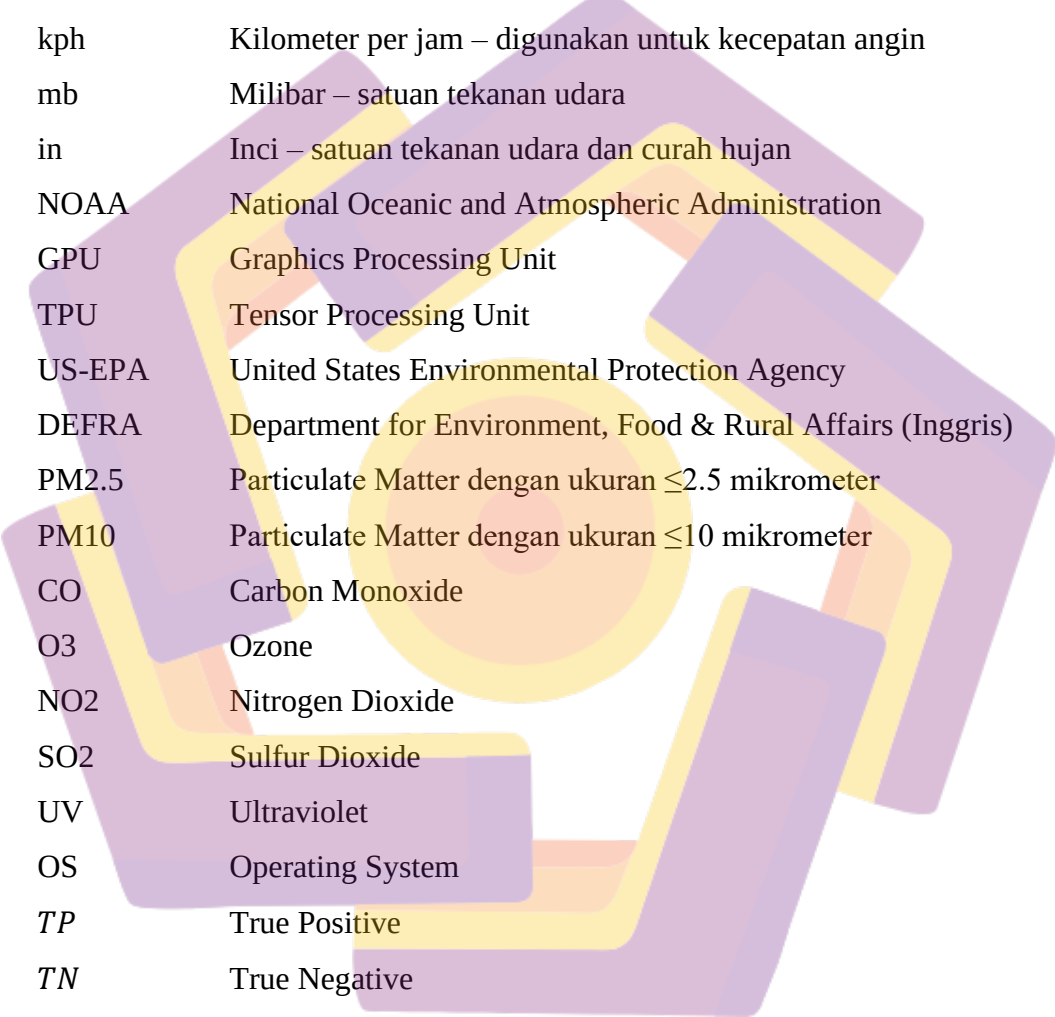


DAFTAR GAMBAR

Gambar 3.1. Alur Penelitian	22
Gambar 3.2. Import Library	23
Gambar 3.3. Setel SEED	24
Gambar 3.4. Memeriksa Nilai Yang Hilang	25
Gambar 3.5. Filter Negara Asia Tenggara	26
Gambar 3.6. Encode Kondisi Cuaca	27
Gambar 3.7. Memilih Fitur	27
Gambar 3.8. Skala Fitur	28
Gambar 3.9. Memecah Data Menjadi Set Pelatihan	29
Gambar 3.10. Membangun Model Bi-LSTM	29
Gambar 3.11. Kompilasi Model	30
Gambar 3.12. Callback Untuk Menghentikan Pelatihan Lebih Awal	30
Gambar 3.13. Melakukan Hyperparameter Tuning	31
Gambar 4.1. Grafik Loss	49
Gambar 4.2. Grafik Accuracy	50
Gambar 4.3. Confusion Matrix	51
Gambar 4.4. Grafik Actual vs Predicted	57

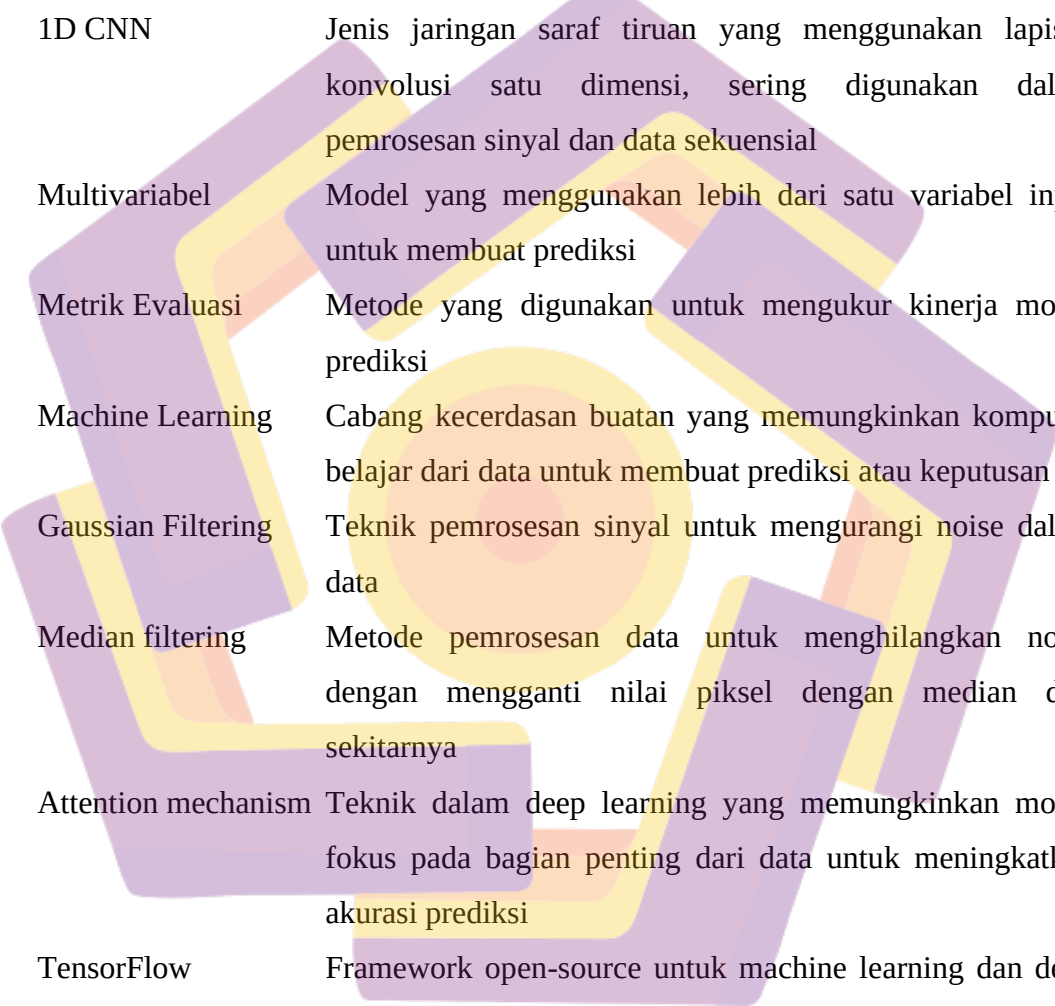
DAFTAR LAMBANG DAN SINGKATAN

Bi-LSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
WRF	Weather Research and Forecasting
TCN	Temporal Convolutional Network
RNN	Recurrent Neural Network
COA	Coyote Optimization Algorithm
ANN	Artificial Neural Network
API	Application Programming Interface
ht	Output hasil penggabungan LSTM forward dan backward
xt	Input data pada waktu t
P	Tekanan udara
ρ	Kerapatan udara
R	Konstanta untuk udara
T	Suhu udara
X'	Data hasil normalisasi
X	Data asli
$Xmin$	Nilai minimum dalam dataset
$Xmax$	Nilai maksimum dalam dataset
y	Prediksi keluaran
W	Bobot jaringan saraf
X	Input data
b	Bias
f	Fungsi aktivasi
n	Jumlah total sampel/data
Σ	Sigma (Penjumlahan)
y_i	Nilai aktual (data sebenarnya) pada sampel ke- i

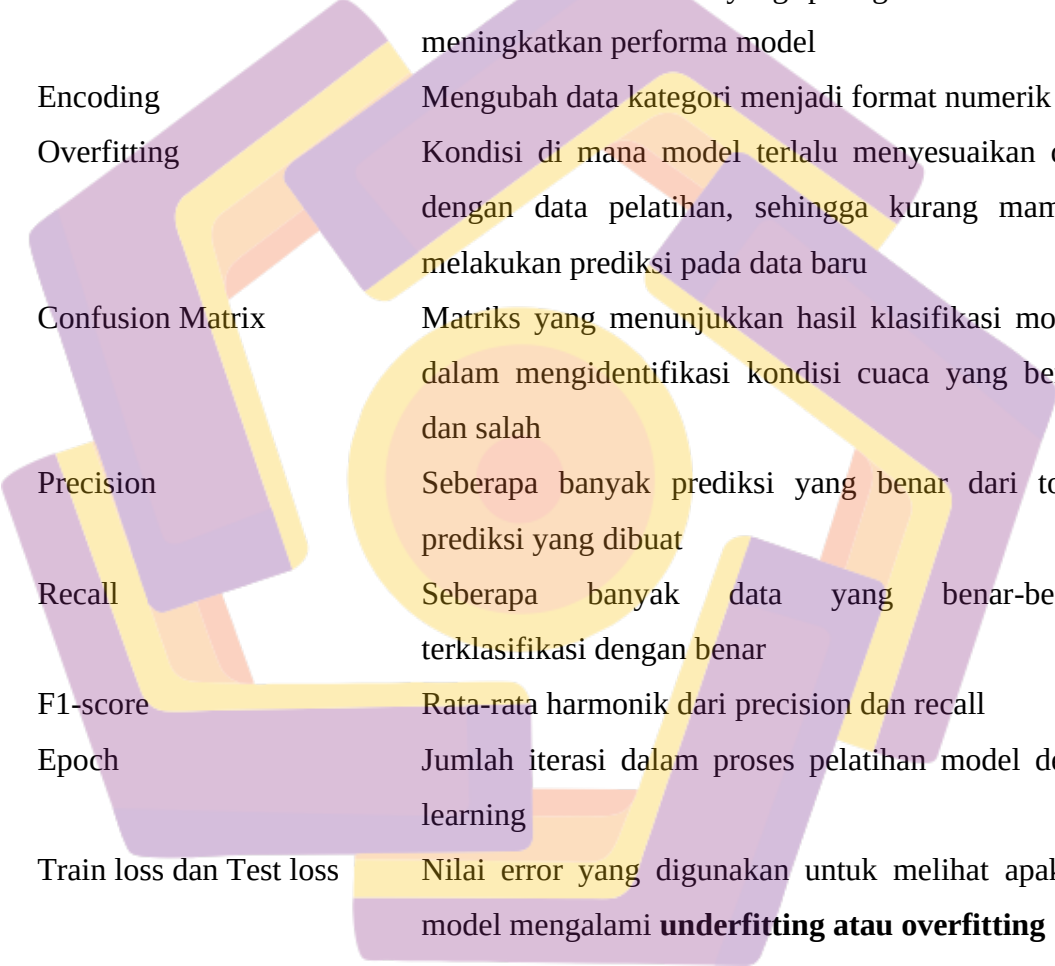


$\hat{y}_i$	Nilai prediksi pada sampel ke- i
°	Derajat
%	Persentase
$\mu\text{g}/\text{m}^3$	Mikrogram per meter kubik – satuan untuk konsentrasi udara
km	Kilometer – digunakan untuk jarak pandang
mph	Mil per jam – digunakan untuk kecepatan angin
kph	Kilometer per jam – digunakan untuk kecepatan angin
mb	Milibar – satuan tekanan udara
in	Inci – satuan tekanan udara dan curah hujan
NOAA	National Oceanic and Atmospheric Administration
GPU	Graphics Processing Unit
TPU	Tensor Processing Unit
US-EPA	United States Environmental Protection Agency
DEFRA	Department for Environment, Food & Rural Affairs (Inggris)
PM _{2.5}	Particulate Matter dengan ukuran ≤ 2.5 mikrometer
PM ₁₀	Particulate Matter dengan ukuran ≤ 10 mikrometer
CO	Carbon Monoxide
O ₃	Ozone
NO ₂	Nitrogen Dioxide
SO ₂	Sulfur Dioxide
UV	Ultraviolet
OS	Operating System
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

DAFTAR ISTILAH



Meteorologi	Ilmu yang mempelajari atmosfer, termasuk cuaca dan iklim
Deep Learning	Subset dari machine learning yang menggunakan jaringan saraf tiruan dengan banyak lapisan untuk menangkap pola kompleks dalam data
1D CNN	Jenis jaringan saraf tiruan yang menggunakan lapisan konvolusi satu dimensi, sering digunakan dalam pemrosesan sinyal dan data sekuensial
Multivariabel	Model yang menggunakan lebih dari satu variabel input untuk membuat prediksi
Metrik Evaluasi	Metode yang digunakan untuk mengukur kinerja model prediksi
Machine Learning	Cabang kecerdasan buatan yang memungkinkan komputer belajar dari data untuk membuat prediksi atau keputusan
Gaussian Filtering	Teknik pemrosesan sinyal untuk mengurangi noise dalam data
Median filtering	Metode pemrosesan data untuk menghilangkan noise dengan mengganti nilai piksel dengan median dari sekitarnya
Attention mechanism	Teknik dalam deep learning yang memungkinkan model fokus pada bagian penting dari data untuk meningkatkan akurasi prediksi
TensorFlow	Framework open-source untuk machine learning dan deep learning
Keras	API tingkat tinggi berbasis TensorFlow yang memudahkan pengembangan model deep learning
Pandas	Library Python yang digunakan untuk manipulasi dan analisis data
NumPy	Library Python yang digunakan untuk komputasi numerik
Matplotlib	Library Python untuk visualisasi data



Seaborn	Library visualisasi data berbasis Matplotlib dengan tampilan lebih menarik
Feature engineering	Proses menciptakan fitur baru atau mengubah fitur yang ada untuk meningkatkan performa model
Hyperparameter tuning	Proses mengoptimalkan parameter model untuk meningkatkan performanya
Feature selection	Proses memilih fitur yang paling relevan untuk meningkatkan performa model
Encoding	Mengubah data kategori menjadi format numerik
Overfitting	Kondisi di mana model terlalu menyesuaikan diri dengan data pelatihan, sehingga kurang mampu melakukan prediksi pada data baru
Confusion Matrix	Matriks yang menunjukkan hasil klasifikasi model dalam mengidentifikasi kondisi cuaca yang benar dan salah
Precision	Seberapa banyak prediksi yang benar dari total prediksi yang dibuat
Recall	Seberapa banyak data yang benar-benar terklasifikasi dengan benar
F1-score	Rata-rata harmonik dari precision dan recall
Epoch	Jumlah iterasi dalam proses pelatihan model deep learning
Train loss dan Test loss	Nilai error yang digunakan untuk melihat apakah model mengalami underfitting atau overfitting

INTISARI

Prediksi cuaca yang buruk sangat berdampak pada banyak aspek kehidupan, seperti perencanaan pertanian, penanggulangan bencana, dan kegiatan sehari-hari masyarakat. Dengan perkembangan kecerdasan buatan, deep learning khususnya Bidirectional Long Short-Term Memory (Bi-LSTM) menawarkan solusi dalam meningkatkan akurasi prediksi cuaca berbasis data historis.

Penelitian ini menggunakan metode eksperimen dengan menerapkan algoritma Bi-LSTM pada dataset historis cuaca dari Kaggle. Langkah-langkah penelitian mencakup akuisisi data, preprocessing (filtering, encoding, normalisasi), serta pembagian data menjadi training (80%) dan testing (20%). Model dilatih menggunakan TensorFlow dan Keras, dengan tuning hyperparameter meliputi dropout rate, jumlah unit LSTM, dan learning rate menggunakan optimizer Adam. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score.

Hasil penelitian menunjukkan bahwa setelah tuning hyperparameter, performa model mengalami peningkatan. Akurasi pada training set meningkat dari 81.27% menjadi 82.56%, sedangkan pada testing set meningkat dari 77.51% menjadi 77.99%. F1-score pada training set meningkat dari 0.7914 menjadi 0.8080, sementara pada testing set naik dari 0.7504 menjadi 0.7647. Precision pada training set meningkat dari 0.8126 menjadi 0.8242, sementara pada testing set meningkat dari 0.7859 menjadi 0.7867. Recall pada training set meningkat dari 0.8127 menjadi 0.8256, sedangkan pada testing set naik dari 0.7751 menjadi 0.7799.

Model Bi-LSTM menunjukkan performa optimal dalam mengklasifikasikan kondisi cuaca yang umum, meskipun masih mengalami kesulitan dalam mengenali kelas cuaca yang jarang terjadi.

Kata kunci: Bi-LSTM, Prediksi Cuaca, Deep Learning, Data Historis, Optimasi Model.

ABSTRACT

Poor weather predictions have a significant impact on many aspects of life, such as agricultural planning, disaster management, and daily community activities. With the advancement of artificial intelligence, deep learning, particularly Bidirectional Long Short-Term Memory (Bi-LSTM), offers a solution to improving the accuracy of weather predictions based on historical data.

This study employs an experimental method by implementing the Bi-LSTM algorithm on a historical weather dataset from Kaggle. The research steps include data acquisition, preprocessing (filtering, encoding, normalization), and splitting the data into training (80%) and testing (20%). The model is trained using TensorFlow and Keras, with hyperparameter tuning involving dropout rate, the number of LSTM units, and learning rate using the Adam optimizer. The model evaluation is conducted using accuracy, precision, recall, and F1-score metrics.

The research results show that after hyperparameter tuning, the model's performance improved. The accuracy on the training set increased from 81.27% to 82.56%, while on the testing set, it increased from 77.51% to 77.99%. The F1-score on the training set improved from 0.7914 to 0.8080, while on the testing set, it rose from 0.7504 to 0.7647. Precision on the training set increased from 0.8126 to 0.8242, while on the testing set, it improved from 0.7859 to 0.7867. Recall on the training set increased from 0.8127 to 0.8256, whereas on the testing set, it rose from 0.7751 to 0.7799.

The Bi-LSTM model demonstrates optimal performance in classifying common weather conditions, although it still struggles to recognize rare weather classes.

Keyword: Bi-LSTM, Weather Prediction, Deep Learning, Historical Data, Model Optimization.