

# **PERBANDINGAN PERFORMA *LONG SHORT-TERM MEMORY* DAN XGBOOST DALAM MEMPREDIKSI CURAH HUJAN HARIAN DI SUMATERA SELATAN**

**OKSI AL HADI**



**DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2024**

@Hak cipta milik IPB University

IPB University



**IPB University**  
Bogor Indonesia

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
    - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
  2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



### @Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

## PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Perbandingan Performa *Long Short-Term Memory* dan *XGBoost* dalam Memprediksi Curah Hujan Harian di Sumatera Selatan” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2024

Oksi Al Hadi  
G1401201045

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## ABSTRAK

OKSI AL HADI. Perbandingan Performa *Long Short-Term Memory* dan XGBoost dalam Memprediksi Curah Hujan Harian di Sumatera Selatan. Dibimbing oleh KUSMAN SADIK dan MOHAMMAD MASJKUR.

Bencana di Indonesia didominasi oleh bencana hidrometeorologi seperti banjir, tanah longsor, dan kekeringan. Salah satu faktor penyebab bencana hidrometeorologi adalah curah hujan. Tercatat telah terjadi 60 kasus bencana hidrometeorologi di Sumatera Selatan yaitu banjir sebanyak 43 kali dan tanah longsor sebanyak 17 kali. Terdapat berbagai algoritme *machine learning* yang memiliki kompleksitas tinggi dan mampu menghasilkan akurasi terbaik pada banyak kasus seperti *long short-term memory* dan XGBoost. Penelitian ini bertujuan untuk membandingkan performa algoritme *long short-term memory* dan XGBoost serta menentukan peubah-peubah yang berpengaruh terhadap curah hujan harian melalui metode *shapley additive explanations*. Dataset yang digunakan merupakan data harian di Provinsi Sumatera Selatan periode 2011 hingga 2020. Peubah yang digunakan ada enam yaitu curah hujan sebagai peubah respon dan lima peubah penjelas lainnya yaitu temperatur, kelembapan udara, kecepatan angin, tekanan udara, dan titik embun. Dalam membandingkan kedua model tersebut, XGBoost memiliki performa yang lebih unggul dibandingkan dengan LSTM dalam memprediksi curah hujan dengan nilai RMSE dan MAE pada hasil prediksi sebesar 5,21 dan 1,93. Peubah penting yang dihasilkan pada pemodelan XGBoost adalah rataan dua data curah hujan sebelumnya dan curah hujan satu hari sebelumnya.

Kata kunci: curah hujan, hidrometeorologi, *long short-term memory*, xgboost

## ABSTRACT

OKSI AL HADI. Performance Comparison of Long Short-Term Memory and XGBoost in Predicting Daily Rainfall in South Sumatra. Supervised by KUSMAN SADIK and MOHAMMAD MASJKUR.

Disasters in Indonesia are dominated by hydrometeorological disasters such as floods, landslides and droughts. One of the factors causing hydrometeorological disasters is rainfall. There have been 60 cases of hydrometeorological disasters in South Sumatra, namely floods 43 times and landslides 17 times. There are various machine learning algorithms that have high complexity and are able to produce the best accuracy in many cases such as long short-term memory and XGBoost. This study aims to compare the performance of long short-term memory and XGBoost algorithms and determine the variables that affect daily rainfall through the shapley additive explanations method. The dataset used is daily data in South Sumatra Province for the period 2011 to 2020. There are six variables used, namely rainfall as the response variable and five other explanatory variables, namely temperature, air humidity, wind speed, air pressure, and dew point. In comparing the two models, XGBoost has superior performance compared to LSTM in predicting rainfall with RMSE and MAE values in the prediction results of 5,21 and 1,93. The important variables generated in XGBoost modelling are the average of the two previous rainfall data and the rainfall one day before.

**Keywords:** hidrometeorologi, long short-term memory, rainfall, xgboost



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024<sup>1</sup>  
Hak Cipta dilindungi Undang-Undang

*Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.*

*Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.*



# **PERBANDINGAN PERFORMA *LONG SHORT-TERM MEMORY* DAN XGBOOST DALAM MEMPREDIKSI CURAH HUJAN HARIAN DI SUMATERA SELATAN**

**OKSI AL HADI**

Skripsi  
sebagai salah satu syarat untuk memperoleh gelar  
Sarjana pada  
Program Studi Statistika dan Sains Data

**DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2024**



**@Hak cipta milik IPB University**

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:  
**Dr. Septian Rahardiantoro, S.Si., M.Si**





Judul Skripsi : Perbandingan Performa *Long Short-Term Memory* dengan XGBoost dalam Memprediksi Curah Hujan Harian di Sumatera Selatan

Nama : Oksi Al Hadi  
NIM : G1401201045

Pembimbing 1:  
Dr. Kusman Sadik, S.Si, M.Si.

Pembimbing 2:  
Ir. Mohammad Masjkur, M.S.

Ketua Departemen Statistika:  
Dr. Bagus Sartono, S.Si., M.Si.  
NIP 197804112005011002

Disetujui oleh



Diketahui oleh

\_\_\_\_\_

Tanggal Ujian:  
22 Juli 2024

Tanggal Lulus:

Hak Cipta Dilindungi Undang-undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah  
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



## PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2023 sampai bulan Juni 2024 ini ialah perbandingan *Statistical machine learning* dengan judul "Perbandingan Performa *Long Short-Term Memory* dan XGBoost dalam Memprediksi Curah Hujan Harian di Sumatera Selatan".

Penulisan karya ilmiah ini tidak luput atas bantuan dan dukungan berbagai pihak sehingga karya ilmiah ini dapat terselesaikan dengan baik. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Almarhum bapa Ferry, mama Kholiyah, kakak Candra, ayuk Fera, adik Fika dan keluarga besar penulis yang selalu memberikan doa, dukungan, dan kasih sayang kepada penulis;
2. Bapak Dr. Kusman Sadik, S.Si., M.Si. dan Bapak Ir. Mohammad Masjkur, M.S. selaku komisi pembimbing yang telah membantu, membimbing, mengarahkan, dan mendukung selama penyusunan karya ilmiah ini;
3. Bapak Ir. Aam Alamudi, M.Si selaku moderator kolokium juga seminar dan Bapak Dr. Septian Rahardianto, S.Si., M.Si selaku penguji luar komisi pembimbing, atas sudut pandang yang diberikan, serta saran dan masukannya dalam rangka perbaikan karya ilmiah;
4. Seluruh dosen dan tenaga pendidik Departemen Statistika yang telah memberi ilmu yang bermanfaat dan menunjang segala kebutuhan penulis selama perkuliahan dan penyusunan karya ilmiah ini;
5. Teman-teman seperjuangan Nachnur, Herdian, Mafazah, Nana, Mercy, Salsa dan Dzakwan yang senantiasa mendukung dan membersamai penulis dalam penyusunan karya ilmiah ini;
6. Teman-teman Treazur/Statistika IPB 57 yang selalu memberikan semangat kepada penulis untuk menyelesaikan karya ilmiah ini; dan
7. Seluruh pihak yang telah membantu penulis dalam penyusunan karya ilmiah ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dalam penulisan karya ilmiah ini masih banyak kekurangan. Penulis berharap karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan..

Bogor, Agustus 2024

*Oksi Al Hadi*

## DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Curah Hujan	3
2.2 Pencilan	3
2.3 <i>Temporal Feature Extraction</i>	4
2.4 <i>Machine Learning</i>	6
2.5 <i>Long Short-Term Memory</i>	6
2.6 <i>Multivariate Long Short-Term Memory</i>	8
2.7 <i>Ensemble Learning</i>	8
2.8 <i>Extreme Gradient Boosting</i>	9
2.9 <i>Walk Forward Validation</i>	10
2.10 <i>Hyperparameter Tuning</i>	11
2.11 Evaluasi Kinerja Model	12
2.12 <i>Shapley Additive Explanation</i>	12
III METODE	14
3.1 Bahan dan Data	14
3.2 Prosedur Analisis Data	14
IV HASIL DAN PEMBAHASAN	17
4.1 Eksplorasi Data	17
4.2 <i>Temporal Feature Extraction</i>	18
4.3 Hasil Sampel <i>Walk Forward Validation</i>	22
4.4 Evaluasi Model LSTM Hasil <i>Walk Forward Validation</i>	24
4.5 Evaluasi Model XGBoost Hasil <i>Walk Forward Validation</i>	24
4.6 Pemodelan LSTM	25
4.7 Pemodelan XGBoost	26
4.8 Perbandingan Performa Model LSTM dan XGBoost	27
4.9 <i>Feature Importance Shapley Additive Explanations</i>	29
V SIMPULAN DAN SARAN	31
5.1 Simpulan	31
5.2 Saran	31
DAFTAR PUSTAKA	32
LAMPIRAN	36
RIWAYAT HIDUP	42



## DAFTAR TABEL

1	Ilustrasi peubah lag	4
2	Ilustrasi peubah <i>rolling mean window</i>	5
3	Peubah-peubah yang digunakan dalam penelitian	14
4	Skema <i>walk forward validation</i>	16
5	Kombinasi <i>hyperparameter tuning</i>	16
6	Statistik deskriptif data curah hujan harian	17
7	Peubah <i>rolling mean window</i>	22
8	Evaluasi terbaik LSTM	24
9	Evaluasi terbaik XGBoost	25
10	Evaluasi LSTM dan XGBoost pada data latih	27
11	Evaluasi LSTM dan XGBoost pada data uji	27

## DAFTAR GAMBAR

1	Arsitektur LSTM (Yudistira 2020)	6
2	Arsitektur umum M-LSTM	8
3	Cara kerja XGBoost (Guo <i>et al.</i> 2020)	9
4	Ilustrasi <i>walk forward validation</i> (Khalid <i>et al.</i> 2021)	11
5	Ilustrasi umum dari proses analisis data	14
6	Ilustrasi <i>walk forward validation</i> skema tahunan	16
7	<i>Boxplot</i> curah hujan periode 2011-2020	17
8	Plot deret waktu curah hujan periode 2011-2020	18
9	Plot PACF curah hujan	19
10	Plot <i>cross-correlation</i> curah hujan terhadap peubah suhu	19
11	Plot <i>cross-correlation</i> curah hujan terhadap peubah kelembapan	20
12	Plot <i>cross-correlation</i> curah hujan terhadap peubah titik embun	20
13	Plot <i>cross-correlation</i> curah hujan terhadap peubah kecepatan udara	21
14	Plot <i>cross-correlation</i> curah hujan terhadap peubah tekanan udara	21
15	Sampel ke-1 hasil <i>walk forward validation</i>	22
16	Sampel ke-2 hasil <i>walk forward validation</i>	23
17	Sampel ke-3 hasil <i>walk forward validation</i>	23
18	Sampel ke-4 hasil <i>walk forward validation</i>	23
19	Sampel ke-5 hasil <i>walk forward validation</i>	23
20	Hasil prediksi metode LSTM pada data uji	25
21	Hasil prediksi metode XGBoost pada data uji	26
22	Perbandingan performa model LSTM dan XGBoost pada data uji	28
23	Perbandingan performa model LSTM dan XGBoost pada data uji	28
24	Skor kepentingan peubah dari model XGBoost menggunakan SHAP menurut interpretasi global	29
25	Skor kepentingan peubah dari model XGBoost menggunakan SHAP menurut interpretasi lokal	30



## DAFTAR LAMPIRAN

1	Lampiran 1 Evaluasi nilai rata-rata hasil <i>walk forward validation</i> dari model LSTM	36
26	Lampiran 2 Evaluasi nilai rata-rata hasil <i>walk forward validation</i> dari model XGBoost	37
27	Lampiran 3 Syntax yang digunakan dalam analisis	38

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
  - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.