

PERBANDINGAN PERFORMA LSTM, BI-LSTM, DAN CNN-LSTM PADA PEMODELAN *STATISTICAL DOWNSCALING* UNTUK DATA CURAH HUJAN DI JAWA BARAT

DINDA KHAMILA NURFATIMAH



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@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 LSTM, Bi-LSTM, dan CNN-LSTM pada Pemodelan *Statistical Downscaling* untuk Data Curah Hujan di Jawa Barat” 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 2025

Dinda Khamila Nurfatihah
G1401211035



ABSTRAK

DINDA KHAMILA NURFATIMAH. Perbandingan Performa LSTM, Bi-LSTM, dan CNN-LSTM pada Pemodelan *Statistical Downscaling* untuk Data Curah Hujan di Jawa Barat. Dibimbing oleh KUSMAN SADIK dan LAILY NISSA ATUL MUALIFAH.

Indonesia merupakan negara kepulauan terbesar di dunia yang terletak di antara dua benua dan dua Samudra menjadikan Indonesia memiliki variabilitas iklim yang sangat tinggi, terutama dalam hal curah hujan. Curah hujan berperan penting dalam berbagai aspek kehidupan manusia. Oleh karena itu, prediksi curah hujan secara akurat penting dilakukan guna menunjang ketahanan pangan dan pembangunan pertanian yang berkelanjutan. Pemodelan untuk memprediksi curah hujan telah banyak dilakukan dengan memanfaatkan data yang dihasilkan oleh *General Circulation Model* (GCM). Namun, GCM umumnya memiliki resolusi spasial yang terlalu kasar untuk diaplikasikan pada skala lokal sehingga pemanfaatannya memerlukan teknik tertentu, yaitu *Statistical Downscaling* (SD). Penelitian ini bertujuan untuk mengembangkan pemodelan SD pada data curah hujan di Jawa Barat dengan menggunakan tiga metode *deep learning*, yaitu *Long Short-Term Memory* (LSTM), *Bidirectional LSTM* (Bi-LSTM), dan *Convolutional Neural Network-LSTM* (CNN-LSTM). Penelitian ini juga bertujuan untuk membandingkan kinerja ketiga model tersebut guna mengidentifikasi pendekatan yang paling sesuai. Dataset yang digunakan mencakup data luaran GCM sebagai peubah prediktor dan data curah hujan lokal dari Badan Meteorologi Klimatologi dan Geofisika (BMKG) sebagai peubah respon pada Stasiun Subang, Bogor, Bandung, dan Majalengka, dengan periode tahun 1995 hingga 2024. Model CNN-LSTM menunjukkan kinerja prediksi terbaik di seluruh stasiun. Model Bi-LSTM menunjukkan performa yang kompetitif dengan nilai RMSE dan MAE yang sedikit lebih tinggi dibandingkan dengan model CNN-LSTM di keempat stasiun. Sebaliknya, Model LSTM menghasilkan nilai RMSE dan MAE yang lebih tinggi, sehingga kurang optimal dibandingkan dua model lainnya.

Kata kunci: Bi-LSTM, CNN-LSTM, curah hujan, LSTM, *statistical downscaling*.

ABSTRACT

DINDA KHAMILA NURFATIMAH. Comparative Performance of LSTM, Bi-LSTM, and CNN-LSTM in Statistical Downscaling Modeling of Rainfall Data in West Java. Supervised by KUSMAN SADIK and LAILY NISSA ATUL MUALIFAH.

Indonesia is the world's largest archipelagic country, located between two continents and two oceans, which results in a high degree of climate variability, particularly in terms of rainfall. Rainfall plays a vital role in various aspects of human life. Therefore, accurate rainfall prediction is essential to support food security and sustainable agriculture development. Rainfall prediction modeling has been widely carried out using data generated by General Circulation Models (GCMs). However, GCMs typically have spatial resolutions that are too coarse for local-scale applications, necessitating the use of techniques such as Statistical Downscaling (SD). This study aims to develop SD models for rainfall prediction in West Java using three deep learning methods: Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and Convolutional Neural Network–LSTM (CNN-LSTM). It also seeks to compare the performance of these models to identify the most suitable approach. The dataset includes GCM outputs as predictor variables and local rainfall data from the Meteorology, Climatology, and Geophysical Agency (BMKG) as the response variable at Subang, Bogor, Bandung, and Majalengka Stations, for the period 1995 to 2024. The CNN-LSTM model demonstrated the best predictive performance across all stations. The Bi-LSTM model also showed competitive performance, with slightly higher RMSE and MAE values compared to the CNN-LSTM model at all four stations. In contrast, the LSTM model yielded higher RMSE and MAE values, making it less optimal than the other two models.

Keywords: Bi-LSTM, CNN-LSTM, LSTM, rainfall, statistical downscaling.



@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.

© Hak Cipta milik IPB, tahun 2025
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 LSTM, BI-LSTM, DAN CNN-LSTM PADA PEMODELAN *STATISTICAL DOWNSCALING* UNTUK DATA CURAH HUJAN DI JAWA BARAT

DINDA KHAMILA NURFATIMAH

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

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@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.



Judul Skripsi : Perbandingan Performa LSTM, Bi-LSTM, dan CNN-LSTM pada
Pemodelan *Statistical Downscaling* untuk Data Curah Hujan di
Jawa Barat

Nama : Dinda Khamila Nurfatimah
NIM : G1401211035

Disetujui oleh

Pembimbing 1:

Dr. Kusman Sadik, S.Si., M.Si.

Pembimbing 2:

Laily Nissa Atul Mualifah, M.Si.

Diketahui oleh

Ketua Program Studi Statistika dan Sains Data:

Dr. Bagus Sartono, S.Si., M.Si.

NIP 197804112005011002

Tanggal Ujian:
21 Juli 2025

Tanggal Lulus:



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 November 2024 sampai bulan Agustus 2025 ini ialah perbandingan *statistical deep learning*, dengan judul "Perbandingan Performa LSTM, Bi-LSTM, dan CNN-LSTM pada Pemodelan *Statistical Downscaling* untuk Data Curah Hujan di Jawa Barat".

Skripsi ini bukan hanya hasil dari usaha akademik semata, tetapi juga perjalanan panjang yang dipenuhi do'a, dukungan, dan semangat dari banyak pihak. Penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua dan keluarga penulis, Nana Rukmana, Pratiwi Suryani, serta Dita, Fina, dan Dila, atas kasih sayang tanpa batas, do'a yang tak pernah putus, dan semangat yang senantiasa menguatkan dalam setiap langkah.
2. Bapak Dr. Kusman Sadik, S.Si., M.Si. dan Ibu Laily Nissa Atul Muallifah, M.Si. selaku dosen pembimbing yang telah membantu, membimbing, mengarahkan, dan mendukung selama penyusunan karya ilmiah ini.
3. Bapak Dr. Bagus Sartono, S.Si., M.Si. dan Aulia Rizki Firdawanti, M.Si. selaku dosen penguji luar dan moderator seminar hasil yang telah memberikan kritik yang membangun dan masukan yang berarti dalam penulisan karya ilmiah ini.
4. Seluruh dosen dan tenaga pendidik Program Studi Statistika dan Sains Data yang telah memberikan ilmu yang bermanfaat serta menunjang segala kebutuhan penulis selama perkuliahan dan penyusunan karya ilmiah ini.
5. Teman-teman Statistika 58, terutama Anis, Nadila, Kamilah, dan Aida, atas kebersamaan, diskusi, dan tawa yang menjadi bagian berharga dalam perjalanan ini.
6. Seluruh pihak yang telah membantu penulis dalam penyusunan karya ilmiah ini yang tidak dapat disebutkan satu-persatu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis dengan rendah hati menerima kritik dan saran yang membangun demi perbaikan di masa mendatang. Penulis berharap skripsi ini dapat memberikan kontribusi positif, baik bagi pengembangan ilmu pengetahuan maupun sebagai referensi untuk penelitian selanjutnya.

Bogor, Agustus 2025

Dinda Khamila Nurfatimah



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.

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Curah Hujan	3
2.2 <i>General Circulation Model</i>	3
2.3 <i>Statistical Downscaling</i>	4
2.4 <i>Deep Learning</i>	5
2.5 <i>Long Short-Term Memory</i>	5
2.6 <i>Bidirectional LSTM</i>	7
2.7 <i>Convolutional Neural Network</i>	8
2.8 CNN-LSTM	10
2.9 <i>K-Nearest Neighbor Imputation</i>	10
2.10 Normalisasi Data	11
2.11 <i>Hyperparameter Tuning</i>	12
2.12 <i>Walk Forward Validation</i>	13
2.13 Evaluasi Kinerja Model	14
III METODE	16
3.1 Data	16
3.2 Prosedur Analisis Data	17
IV HASIL DAN PEMBAHASAN	20
4.1 Praproses Data	20
4.2 Eksplorasi Data	20
4.3 Normalisasi Data	22
4.4 Transformasi Format Data	23
4.5 Evaluasi Model Hasil <i>Walk Forward Validation</i>	23
4.6 Pemodelan LSTM, Bi-LSTM, dan CNN-LSTM	25
4.7 Perbandingan Performa LSTM, Bi-LSTM, dan CNN-LSTM	30
V SIMPULAN DAN SARAN	33
5.1 Simpulan	33
5.2 Saran	33
DAFTAR PUSTAKA	34
LAMPIRAN	39
RIWAYAT HIDUP	51



DAFTAR TABEL

1	Deskripsi stasiun curah hujan yang digunakan pada penelitian	16
2	Kombinasi <i>hyperparameter</i> untuk proses <i>tuning</i>	18
3	Statistik deskriptif curah hujan keempat stasiun	21
4	Hasil normalisasi data	23
5	Kombinasi <i>hyperparameter</i> terbaik keempat stasiun	24
6	Konfigurasi <i>hyperparameter</i>	26
7	Evaluasi LSTM, Bi-LSTM, dan CNN-LSTM pada data latih	30
8	Evaluasi LSTM, Bi-LSTM, dan CNN-LSTM pada data uji	32

DAFTAR GAMBAR

1	Ilustrasi proses SD (Sutikno 2008)	4
2	Arsitektur LSTM (Goyal <i>et al.</i> 2021)	6
3	Arsitektur Bi-LSTM (Bohara <i>et al.</i> 2022)	8
4	Arsitektur CNN (Ahmad <i>et al.</i> 2023)	9
5	Arsitektur CNN-LSTM (Widiputra 2021)	10
6	Ilustrasi <i>walk forward validation</i> (dimodifikasi dari Mahmood dan Avşar (2021))	14
7	Ilustrasi <i>grid</i> model GCM	16
8	Skema jendela <i>walk forward validation</i> pada Stasiun Subang	18
9	Contoh hasil imputasi nilai hilang	20
10	<i>Boxplot</i> curah hujan (a) Stasiun Subang, (b) Stasiun Bogor, (c) Stasiun Bandung, dan (d) Stasiun Majalengka	21
11	(a) Plot dekomposisi data curah hujan Stasiun Subang, (b) komponen tren, (c) komponen musiman	22
12	<i>Boxplot</i> RMSE hasil kombinasi <i>hyperparameter</i> terbaik untuk setiap skenario pergeseran jendela (a) Stasiun Subang, (b) Stasiun Bogor, (c) Stasiun Bandung, (d) Stasiun Majalengka	25
13	plot <i>train loss</i> dan <i>validation loss</i> pada model LSTM, Bi-LSTM, dan CNN-LSTM di Stasiun Subang	27
14	Plot hasil evaluasi model LSTM, Bi-LSTM, dan CNN-LSTM terhadap data latih pada Stasiun Bogor	28
15	Plot hasil evaluasi model LSTM, Bi-LSTM, dan CNN-LSTM terhadap data latih pada Stasiun Bogor dengan <i>epoch</i> sebesar 50	29
16	Hasil prediksi LSTM, Bi-LSTM, dan CNN-LSTM pada data uji (a) Stasiun Subang, (b) Stasiun Bogor, (c) Stasiun Bandung, dan (d) Stasiun Majalengka	31

DAFTAR LAMPIRAN

1	Lampiran 1 Skema jendela <i>walk forward validation</i>	39
2	Lampiran 2 Plot dekomposisi deret waktu	40



- 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.

3	Lampiran 3 Evaluasi nilai rata-rata RMSE hasil <i>hyperparameter tuning</i>	42
4	Lampiran 4 Plot <i>train loss</i> vs <i>validation loss</i> pada model LSTM, Bi-LSTM, dan CNN-LSTM	45
5	Lampiran 5 Plot hasil evaluasi model LSTM, Bi-LSTM, dan CNN-LSTM terhadap data latih	47