

**PREDIKSI CURAH HUJAN BERPOTENSI BANJIR
MENGUNAKAN ALGORITMA *LONG SHORT-TERM
MEMORY* (LSTM) DAN *ISOLATION FOREST***

**Tesis
untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-2 Program Studi
Magister Sistem Informasi**



**Ahmad Wijayanto
30000322140013**

**SEKOLAH PASCASARJANA
UNIVERSITAS DIPONEGORO
SEMARANG
2024**

HALAMAN PENGESAHAN

TESIS :

**PREDIKSI CURAH HUJAN BERPOTENSI BANJIR MENGGUNAKAN
ALGORITMA *LONG SHORT TERM MEMORY* (LSTM)
DAN *ISOLATION FOREST***

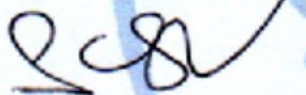
Oleh:

Ahmad Wijayanto
30000322410013

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 05 Agustus 2024 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Semarang, 05 Agustus 2024
Mengetahui,

Pembimbing I



Dr. Aris Sugiharto, S.Si., M.Kom.
NIP 197108111997021004

Pembimbing II



Dr. Drs. Rukun Santoso, M.Si.
NIP 196502251992011001

Penguji I



Vincensius Ghanawan Slamet Kadarisman,
S.Si., M.Si., Ph.D.
NIP 197105221997021001

Penguji II



Dr. Eng. Ir. Wahyul Amien Syafei, S.T.,
M.T., IPM
NIP 197112181995121001

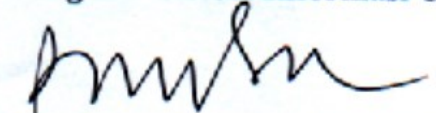
Mengetahui :

**Dekan Sekolah Pascasarjana
Universitas Diponegoro**

**Ketua Program Studi
Magister Sistem Informasi UNDIP**



Prof. Ir. M. Agung Wibowo, M.M.,
M.Sc., Ph.D.
NIP 196702081994031005

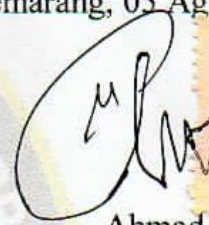


Drs. Bayu Surarso, M.Sc., Ph.D.
NIP 196311051988031001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam tesis ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Semarang, 05 Agustus 2024



Ahmad Wijayanto



SEKOLAH PASCASARJANA

**PERNYATAAN PERSETUJUAN PUBLIKASI TESIS
UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini :

Nama : Ahmad Wijayanto
NIM : 30000322410013
Program Studi : Magister Sistem Informasi
Program : Sekolah Pascasarjana
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif atas karya ilmiah saya yang berjudul :

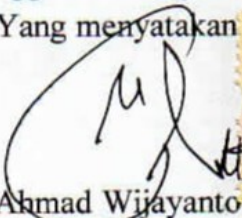
**PREDIKSI CURAH HUJAN BERPOTENSI BANJIR MENGGUNAKAN
ALGORITMA *LONG SHORT-TERM MEMORY* (LSTM)
DAN *ISOLATION FOREST***

beserta perangkat yang ada. Dengan Hak bebas Royalti Noneksklusif ini Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database) merawat, dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Dibuat di : Semarang

Pada tanggal : 05 Agustus 2024

Yang menyatakan


Ahmad Wijayanto

NIM 30000322410013



KATA PENGANTAR

Segala Puji dan syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmatNya sehingga penulis dapat menyelesaikan tesis dengan judul “Prediksi Curah Hujan Berpotensi Banjir Menggunakan Algoritma *Long Short-Term Memory* (LSTM) Dan *Isolation Forest*” guna memenuhi sebagian persyaratan untuk memperoleh gelar magister pada program studi Magister Sistem Informasi, Sekolah Pascasarjana Universitas Diponegoro.

Penulis menyadari kelemahan serta keterbatasan yang ada sehingga dalam menyelesaikan tesis ini memperoleh bantuan dari berbagai pihak, dalam kesempatan ini penulis menyampaikan ucapan terima kasih kepada :

1. Bapak Prof. Ir. Mochamad Agung Wibowo, M.M., M.Sc., Ph.D. selaku dekan Sekolah Pascasarjana Universitas Diponegoro.
2. Bapak Drs. Bayu Surarso, M.Sc., Ph.D. selaku Ketua Program Studi Magister Sistem Informasi Universitas Diponegoro.
3. Bapak Dr. Aris Sugiharto, S.Si, M.Kom selaku Dosen Pembimbing I, yang telah memberikan bimbingan, arahan, serta kritik kepada penulis hingga bisa selesai dalam penyusunan tesis.
4. Bapak Dr. Rukun Santoso M.Si. selaku dosen Pembimbing II, yang sudah membimbing, memberikan arahan, dan masukan yang baik kepada penulis sehingga dapat penulis menyelesaikan laporan tesis ini.
5. Bapak Vincensius Gunawan Slamet Kadarisman, S.Si., M.Si., Ph.D selaku Dosen Ketua Penguji yang telah memberikan banyak saran dan masukan yang baik kepada penulis.
6. Bapak Dr. Eng. Ir. Wahyul Amien Syafei, S.T., M.T., IPM sebagai Dosen Penguji yang telah memberikan banyak saran dan masukan dalam penyusunan tesis.
7. Semua pihak yang tidak dapat disebutkan satu persatu, yang telah membantu sampai dengan terselesaikannya tesis ini.

Penulis menyadari bahwa tesis ini masih banyak kekurangan baik isi maupun susunannya. Semoga tesis ini dapat bermanfaat tidak hanya bagi penulis juga bagi para pembaca.

Semarang, 05 Agustus 2024

Ahmad Wijayanto



DAFTAR ISI

PERNYATAAN.....	ii
PERNYATAAN PERSETUJUAN PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	vi
DAFTAR TABEL.....	viii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Tujuan.....	4
1.3 Manfaat.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Penelitian Terdahulu.....	5
2.2 Pengertian Banjir.....	7
2.2.1 Jenis Banjir.....	7
2.2.2 Faktor-Faktor Penyebab Banjir.....	8
2.2.3 Daerah Rawan Banjir.....	9
2.2.4 Curah Hujan Area.....	10
2.3 <i>Long Short Term Memory</i> (LSTM).....	12
2.4 <i>Gaussian Noise</i> Distribusi Normal.....	15
2.5 <i>Isolation Forest</i> (IF).....	17
2.6 Validasi Model.....	19
2.6.1 <i>Mean Squared Error</i> (MSE).....	19
2.6.2 <i>Root Mean Squared Error</i> (RMSE).....	20
BAB III METODE PENELITIAN.....	21
3.1 Bahan dan Alat Penelitian.....	21
3.1.1 Bahan Penelitian.....	21
3.1.2 Alat.....	21
3.2 Prosedur Penelitian.....	21
3.2.1 Pengumpulan Data.....	22
3.2.2 Pra-Pemrosesan Data.....	22
3.2.3 Pemrosesan Data.....	25
3.2.4 Evaluasi dan Eksperimen Model Prediksi.....	26
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	28
4.1 Hasil Penelitian.....	28
4.1 Pembahasan.....	62
BAB V KESIMPULAN PENELITIAN.....	76
5.1 Kesimpulan.....	76
5.2 Saran.....	76
DAFTAR PUSTAKA.....	77

DAFTAR GAMBAR

Gambar 4. 1 Aturan Empiris (Blaine, 2023)	32
Gambar 4. 2 Grafik Data Aktual 2010 s.d 2024	40
Gambar 4. 3 Pola Curah Hujan Harian Data Aktual Selama 1 Bulan Pada Januari 2010	40
Gambar 4. 4 Pola Curah Hujan Harian Selama 3 Bulan dari Januari-Maret 2010	41
Gambar 4. 5 Pola Curah Hujan Harian Selama 1 Tahun dari Januari-Desember 2010	41
Gambar 4. 6 Grafik Data Setelah Augmentasi dan <i>Reshape</i>	42
Gambar 4. 7 Grafik Data Setelah Augmentasi dan <i>Reshape</i> Selama 30 Hari	42
Gambar 4. 8 Grafik Data Setelah Augmentasi dan <i>Reshape</i> Selama 90 Hari	43
Gambar 4. 9 Grafik Data Setelah Augmentasi dan <i>Reshape</i> Selama 365 Hari	43
Gambar 4. 10 Hasil <i>K-Fold Cross Validation</i>	45
Gambar 4. 11 Grafik Komposisi <i>K Fold</i> Secara <i>Random</i>	46
Gambar 4. 12 Hasil Grafik <i>Loss Function</i> tanpa <i>Gaussian Noise</i>	57
Gambar 4. 13 Hasil Grafik <i>Loss Function</i> Menggunakan <i>Gaussian Noise</i>	57
Gambar 4. 14 <i>Loss Function</i> Menggunakan <i>Cross Validation</i>	58
Gambar 4. 15 Hasil Validasi Prediksi tanpa <i>Gaussian Noise</i>	58
Gambar 4. 16 Hasil Validasi Prediksi dengan <i>Gaussian Noise</i>	59
Gambar 4. 17 Hasil <i>Forecasting</i> dan Identifikasi Data Curah Hujan Potensi Banjir Januari 2022	61
Gambar 4. 18 Hasil <i>Forecasting</i> dan Identifikasi Data Curah Hujan Potensi Banjir Februari - April 2024	62
Gambar 4. 19 <i>Flowchart Processing Data</i>	63
Gambar 4. 20 Grafik 30 hari <i>Forecasting</i> Curah Hujan LSTM	65
Gambar 4. 21 Grafik 30 hari <i>Forecasting</i> Curah Hujan LSTM <i>Gaussian Noise</i> ..	65
Gambar 4. 22 Perbandingan <i>Real Value</i> dan <i>Forecasting</i> Data Selama 1 Bulan Pada Januari 2022	66
Gambar 4. 23 Perbandingan <i>Real Value</i> dan <i>Forecasting</i> Data Selama 3 Bulan Pada Februari 2024 s.d April 2024	67
Gambar 4. 24 Hasil Identifikasi <i>Forecasting</i> Menggunakan <i>IF</i> Selama 1 Bulan Pada Bulan Januari 2022	69
Gambar 4. 25 Hasil Identifikasi <i>Forecasting</i> Menggunakan <i>IF</i> Selama 3 Bulan Pada Bulan Februari s.d April Tahun 2024	70
Gambar 4. 26 Hasil Identifikasi <i>Forecasting</i> Menggunakan <i>IF</i> Selama 1 Bulan Juli Tahun 2024	70
Gambar 4. 27 Hasil Identifikasi <i>Forecasting</i> Menggunakan <i>IF</i> Selama 1 Tahun 2024	71
Gambar 4. 28 Hasil <i>Forecasting</i> 30 Hari dibanding Beberapa Tahun Belakang ..	72

Gambar 4. 29 Hasil *Forecasting* 90 Hari dibanding Beberapa Tahun Belakang ..73



SEKOLAH PASCASARJANA

DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu.....	6
Tabel 2. 2 Keadaan curah hujan dan intensitas curah hujan	11
Tabel 3. 1 <i>Dataset</i> Curah Hujan.....	21
Tabel 3. 2 Data Sebelum Augmentasi.....	23
Tabel 4. 1 Hasil Percobaan Perubahan Standard Deviasi	37
Tabel 4. 2 Hasil Data <i>Gaussian Noise</i> Dengan 5 <i>Augment Factor</i>	38
Tabel 4. 3 Hasil <i>Reshape</i> Sampel Data	39
Tabel 4. 4 Hasil Sampel Data <i>Splitting K-Fold Cross Validation</i>	44
Tabel 4. 5 Tabel Jumlah Data Pada Proses <i>K-Fold Cross Validation</i>	46
Tabel 4. 6 Tabel Parameter Terbaik.....	55
Tabel 4. 7 Hasil Percobaan MSE dan RMSE Model	56
Tabel 4. 8 Perbandingan Hasil LSTM dengan Penelitian Terdahulu.....	68
Tabel 4. 9 Perbandingan Hasil Dengan Data Yang Sama.....	68
Tabel 4. 10 Hasil <i>Forecasting</i> Selama 1 Tahun di 2024.....	71
Tabel 4. 11 Data Aktual Bencana.....	74

SEKOLAH PASCASARJANA

Lambang	Arti Lambang
$\bar{R}$	Curah Hujan Area (mm)
n	Ukuran sampel
f_t	Gerbang lupa / <i>forget gate</i>
i_t	Gerbang masukan (<i>input gate</i>)
o_t	Gerbang keluaran (<i>output gate</i>)
c_t	Sel memori (<i>cell state</i>)
h_t	Keadaan tersembunyi (<i>hidden state</i>)
x_t	Data <i>input</i>
W_{ij}	Bobot
h_{t-1}	Keadaan tersembunyi sebelumnya pada sel LSTM
c_{t-1}	Sel memori sebelumnya pada sel LSTM
b_{ij}	Bias
σ	Sigmoid
$\tanh$	Tangen hiperbolik
s	Himpunan kasus
$h(x)$	Jumlah langkah jarak vertikal
$c(n)$	Faktor skala
E	Panjang jalur

Singkatan	Kepanjangan Singkatan
LSTM	<i>Long Short Term Memory</i>
IF	<i>Isolation Forest</i>
MSE	<i>Means Squared Error</i>
RMSE	<i>Root Mean Squared Error</i>
BMKG	Badan Meteorologi Klimatologi dan Geofisika
BPBD	Badan Penanggulangan Bencana Daerah

SEKOLAH PASCASARJANA

**PREDIKSI CURAH HUJAN BERPOTENSI BANJIR MENGGUNAKAN
ALGORITMA *LONG SHORT-TERM MEMORY* (LSTM)
DAN *ISOLATION FOREST***

ABSTRAK

Permasalahan banjir yang sering terjadi di berbagai daerah dapat disebabkan oleh faktor curah hujan yang sulit diprediksi. Penelitian ini bertujuan untuk menghasilkan model prediksi dan peramalan jangka panjang curah hujan menggunakan algoritma *Long Short Term Memory* (LSTM) serta identifikasi pengambilan keputusan curah hujan berpotensi banjir menggunakan *Isolation Forest*. Algoritma LSTM digunakan untuk memprediksi curah hujan dengan hasil *training* dan *validation Mean Squared Error* (MSE) sebesar 17,04 dan *Root Mean Squared Error* (RMSE) sebesar 4,12. Probabilitas curah hujan dalam jangka panjang diperoleh dengan penggunaan *Gaussian Noise* pada LSTM yang dipengaruhi oleh tingkat akurasi terbaik. Selanjutnya, algoritma *Isolation Forest* digunakan untuk mengidentifikasi curah hujan yang berpotensi menyebabkan banjir. Hasil dari prediksi ini dapat digunakan untuk menentukan pengambilan keputusan mengenai kejadian curah hujan berpotensi banjir. Dengan hasil probabilitas curah hujan dalam jangka panjang dan identifikasi dini potensi banjir, dapat dimanfaatkan untuk merancang infrastruktur yang lebih tahan banjir, seperti sistem drainase yang lebih baik, pembangunan fasilitas pengendalian banjir yang sesuai, serta perhitungan kekuatan struktur di daerah rawan banjir.

Kata kunci : prediksi, LSTM, *Isolation Forest*.

SEKOLAH PASCASARJANA

FORECASTING RAINFALL AND FLOOD POTENTIAL USING LONG SHORT-TERM MEMORY (LSTM) AND ISOLATION FOREST ALGORITHMS

ABSTRACT

The recurring flood problems in various regions can be attributed to the unpredictable nature of rainfall patterns. This research aims to develop a long-term rainfall prediction and forecasting model using the Long Short-Term Memory (LSTM) algorithm, as well as identify rainfall decisions with potential flood risks using the Isolation Forest algorithm. The LSTM algorithm was employed to predict rainfall, with the training and validation results yielding a Mean Squared Error (MSE) of 17.04 and a Root Mean Squared Error (RMSE) of 4.12. The long-term rainfall probability was obtained by incorporating Gaussian Noise into the LSTM, which was influenced by the best accuracy level. Furthermore, the Isolation Forest algorithm was used to identify rainfall patterns with the potential to cause floods. The results of this prediction can be used to determine decision-making regarding rainfall events with flood potential. By obtaining long-term rainfall probability and early identification of potential flood risks, the information can be utilized to design infrastructure that is more resilient to floods, such as improved drainage systems, the construction of appropriate flood control facilities, and the consideration of structural strength in flood-prone areas.

Keywords : prediksi, LSTM, Isolation Fores

SEKOLAH PASCASARJANA