



**PERBANDINGAN ALGORITMA RANDOM FOREST DAN
DECISION TREE UNTUK PREDIKSI CURAH HUJAN DAN
BANJIR DI JAKARTA PUSAT**

**TUGAS AKHIR
SKRIPSI**

**ZIDANE FAHREZI
41522010008**

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**



**PERBANDINGAN ALGORITMA RANDOM FOREST DAN
DECISION TREE UNTUK PREDIKSI CURAH HUJAN DAN
BANJIR DI JAKARTA PUSAT**

**TUGAS AKHIR
SKRIPSI**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

ZIDANE FAHREZI
41522010008
**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Zidane Fahrezi

NIM : 41522010008

Program Studi : Fakultas Ilmu Komputer / Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

"Perbandingan Algoritma Random Forest dan Decision Tree untuk Prediksi Curah Hujan dan Banjir Harian di Jakarta Pusat" adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 28 Januari 2026



Zidane Fahrezi

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name	: Zidane Fahrezi
NIM /Student ID Number	: 41522010008
Program Studi /Study Program	: Teknik Informatika

Dengan Judul Tugas Akhir

/The title:

**“PERBANDINGAN ALGORITMA RANDOM FOREST DAN DECISION TREE
UNTUK PREDIKSI CURAH HUJAN DAN BANJIR DI JAKARTA PUSAT”**

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

13 Februari 2026

dengan nilai persentase sebesar :

/with a percentage value of:

22%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer**
Universitas Mercu Buana. */declared to meet standards in accordance with applicable
regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

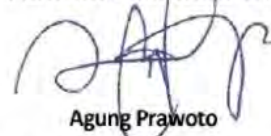
File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/1RabyAODQTX8NAqVrOdZ829pjJ49m6qE/view?usp=drive_link



Jakarta, 13 Februari 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto, S.Kom., B.Sc

NIK : 322970503

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : ZIDANE FAHREZI
NIM : 41522010008
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir : Perbandingan Algoritma Random Forest Dan
Decision Tree Untuk Prediksi Curah Hujan Dan
Banjir Di Jakarta Pusat

Telah berhasil dipertahankan pada sidang tanggal 28 Januari 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing


Ir. Emil Robert Kaburuan, Ph.D., IPM., ASEAN Eng.

NIDN/NUPTK: 0429058004

MERCU BUANA

Jakarta, 28 Januari 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., MTI
NIDN/NUPTK: 0320037002

Ketua Program Studi
Teknik Informatika



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0225067701

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Bapak Ir. Emil Robert Kaburuan, Ph.D., IPM., ASEAN Eng. selaku dosen pembimbing Tugas Akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.
6. Seluruh teman-teman saya yang telah memberikan dukungan, semangat, serta kebersamaan yang berarti selama masa studi.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 30 Desember 2025

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

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

Tugas Akhir ini diajukan oleh:

Nama : Zidane Fahrezi
NIM : 41522010008
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi : Perbandingan Algoritma Random Forest dan Decision Tree untuk Prediksi Curah Hujan dan Banjir Harian di Jakarta Pusat

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 28 Januari 2026

Yang menyatakan,



Zidane Fahrezi

PERBANDINGAN ALGORITMA RANDOM FOREST DAN DECISION TREE UNTUK PREDIKSI CURAH HUJAN DAN BANJIR DI JAKARTA PUSAT

ZIDANE FAHREZI

ABSTRAK

Curah hujan merupakan salah satu faktor utama yang memengaruhi terjadinya banjir di wilayah perkotaan, khususnya di Jakarta Pusat. Untuk mendukung upaya mitigasi banjir, diperlukan pendekatan berbasis *Machine learning* yang mampu melakukan prediksi curah hujan serta klasifikasi potensi banjir secara dini. Penelitian ini bertujuan untuk membahas penerapan dan perbandingan algoritma *Decision Tree* dan *Random Forest* dalam memprediksi curah hujan harian serta mengklasifikasikan kondisi banjir dan tidak banjir. Data yang digunakan berupa data iklim harian wilayah Jakarta Pusat periode 2018–2025. Metodologi penelitian mencakup tahapan *data preprocessing*, *Feature engineering* berbasis waktu, pemanfaatan data historis melalui fitur *lag* dan *rolling window*, pembentukan label banjir berdasarkan ambang batas curah hujan, serta pemodelan menggunakan pendekatan *time series*. Proses evaluasi dilakukan untuk menilai kemampuan model dalam mempelajari pola curah hujan dan kejadian banjir. Penelitian ini diharapkan dapat memberikan kontribusi dalam penerapan *Machine learning* pada analisis data iklim serta menjadi referensi dalam pengembangan sistem prediksi banjir berbasis data historis.

Kata kunci: Curah hujan, Banjir, *Machine learning*, *Random Forest*, *Decision Tree*.

PERBANDINGAN ALGORITMA RANDOM FOREST DAN DECISION TREE UNTUK PREDIKSI CURAH HUJAN DAN BANJIR DI JAKARTA PUSAT

ZIDANE FAHREZI

ABSTRACT

Rainfall is one of the main factors influencing the occurrence of flooding in urban areas, particularly in Jakarta. To support flood mitigation efforts, a Machine learning-based approach is required to enable rainfall prediction and early classification of flood potential. This study aims to discuss the application and comparison of Decision Tree and Random Forest algorithms in predicting daily rainfall and classifying flood and non-flood conditions. The dataset used consists of daily climate data from Central Jakarta and for the period 2018 - 2025. The research methodology includes data preprocessing, time-based Feature engineering, utilization of historical data through lag features and rolling windows, construction of flood labels based on rainfall threshold values, and modeling using a time series approach. The evaluation process is conducted to assess the models' ability to learn rainfall patterns and flood occurrences. This study is expected to contribute to the application of Machine learning in climate data analysis and to serve as a reference for the development of flood prediction systems based on historical data.

Keywords: rainfall, flood, Machine learning, Random Forest, Decision Tree.

DAFTAR ISI

HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu	6
2.2 Teori Utama	11
2.2.1 Banjir sebagai Bencana Hidrometeorologi	11
2.2.2 Curah Hujan dan Karakteristik Data Iklim	11
2.2.3 Data Deret Waktu (Time Series)	12
2.2.4 Data Iklim BMKG sebagai Sumber Dataset.....	12
2.2.5 Pra-pemrosesan Data (Data Preprocessing).....	13
2.2.6 <i>Machine learning</i> dalam Analisis Data Iklim.....	14
2.2.7 Regresi dalam <i>Machine learning</i>	15
2.2.8 Klasifikasi dalam <i>Machine learning</i>	15
2.2.9 Algoritma <i>Decision Tree</i>	16
2.2.10 Algoritma <i>Random Forest</i>	17
2.2.11 <i>Feature engineering</i> pada Data Iklim	18

2.2.12 Perbandingan Model <i>Decision Tree</i> dan <i>Random Forest</i>	18
2.2.13 Evaluasi Model Prediksi dan Klasifikasi	19
2.3 Teori Pendukung	21
2.3.1 Prediksi Curah Hujan Berbasis Data Deret Waktu	21
2.3.2 Klasifikasi Banjir Berdasarkan Curah Hujan dan Variabel Meteorologis 21	
2.3.3 Peran Algoritma <i>Machine learning</i> dalam Prediksi dan Klasifikasi	22
BAB III METODE PENELITIAN	24
3.1 Jenis Penelitian	24
3.2 Tahapan Penelitian	24
3.2.1 Pengumpulan Data	26
3.2.2 Preprocessing Data	26
3.2.3 Feature engineering Berbasis Deret Waktu	26
3.2.4 Pembentukan Label Kejadian Banjir	27
3.2.5 Pembagian Data	27
3.2.6 Pembangunan Model <i>Machine learning</i>	27
3.2.7 Evaluasi Model	28
3.2.8 Perbandingan Model dan Pemilihan Model Terbaik	28
3.2.9 Prediksi Curah Hujan dan Banjir Lima Hari ke Depan	28
3.2.10 Interpretasi Proses Penelitian	29
BAB IV PEMBAHASAN	30
4.1 Pengumpulan Data	30
4.2 Pra-pemrosesan Data	32
4.2.1 Tinjauan Data Awal	32
4.2.2 Pembersihan Nilai Anomali	34
4.2.3 Penanganan Data Hilang (Missing Values)	35
4.3 Pembentukan Label Kejadian Banjir	38
4.3.1 Dasar Penentuan Label Banjir	38
4.3.2 Proses Pelabelan Data	39
4.3.3 Distribusi Kelas dan Ketidakseimbangan Data	40
4.3.4 Pencegahan Kebocoran Data (Data Leakage)	42
4.4 Rekayasa Fitur (Feature engineering)	42
4.4.1 Inisialisasi Dataset untuk Rekayasa Fitur	42
4.4.2 Pembentukan Fitur <i>Lag</i>	43
4.4.3 Pembentukan Fitur Rolling Mean	44
4.4.4 Hasil Akhir Rekayasa Fitur	44
4.5 Pembagian Data	46
4.5.1 Pendekatan Time-Based Split	46
4.5.2 Data Splitting	47
4.5.3 Pembagian Data untuk Regresi dan Klasifikasi	48
4.6 Pemodelan Regresi Curah Hujan	48

4.6.1 Hasil Regresi pada Skenario 70:30	49
4.6.2 Hasil Regresi pada Skenario 80:20	49
4.6.3 Perbandingan Kinerja Model Regresi	50
4.6.4 Model Regresi Terbaik	51
4.7 Pemodelan Klasifikasi Kejadian Banjir	52
4.7.1 Hasil Klasifikasi pada Skenario 70:30	52
4.7.2 Hasil Klasifikasi pada Skenario 80:20	54
4.7.3 Analisis Confusion Matrix dan Perbandingan Model	56
4.7.4 Model Klasifikasi Terbaik	57
4.8 Integrasi Model Regresi dan Klasifikasi pada Prediksi Lima Hari	58
4.8.1 Analisis Hasil Prediksi Curah Hujan Lima Hari	60
BAB V KESIMPULAN DAN SARAN	63
5.1 Kesimpulan	63
5.2 Saran	64
DAFTAR PUSTAKA	65
LAMPIRAN	69



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	6
Tabel 2.2 Confusion Matrix Klasifikasi	21
Tabel 4.1 Data split 70:30 dan 80:20	48
Tabel 4.2 Metrik Evaluasi Model Regresi 70:30	49
Tabel 4.3 Metrik Evaluasi Model Regresi 80:20	50
Tabel 4.4 Evaluasi Model Klasifikasi Banjir 70:30	53
Tabel 4.5 Evaluasi Model Klasifikasi Banjir 80:20	55



DAFTAR GAMBAR

Gambar 3.1 Flowchart Tahapan Penelitian	25
Gambar 4.1 Website “dataonline.bmkg.go.id” Pengumpulan Dataset.....	31
Gambar 4.2 Sampel Data Awal sebelum Pra-pemrosesan	33
Gambar 4.3 Tipe data masing-masing kolom	33
Gambar 4.4 Kode Penanganan Nilai Anomali dan Data Kosong pada Dataset..	34
Gambar 4.5 Contoh Nilai Anomali Curah Hujan sebelum Pembersihan.....	35
Gambar 4.6 Data Curah Hujan setelah Pembersihan Nilai Anomali	35
Gambar 4.7 Kode Tahapan Imputasi Missing Value	36
Gambar 4.8 Contoh Data dengan Missing Values sebelum Imputasi	37
Gambar 4.9 Data setelah Proses Imputasi Missing Values	37
Gambar 4.10 Hubungan Curah Hujan Harian dengan Label Kejadian Banjir	39
Gambar 4.11 Contoh Dataset dengan Kolom Label Kejadian Banjir	40
Gambar 4.12 Distribusi Kelas Kejadian Banjir dan Tidak Banjir.....	41
Gambar 4.13 Persentase Distribusi Kejadian Banjir	41
Gambar 4.14 Inisialisasi Dataset untuk Feature engineering	43
Gambar 4.15 Lag Features (1, 3, 7 hari).....	43
Gambar 4.16 Kode Rolling Mean (3 & 7 hari)	44
Gambar 4.17 Hasil Akhir Feature engineering.....	45
Gambar 4.18 Data split 70:30 dan 80:20.....	47
Gambar 4.19 Metrik Evaluasi Regresi 70:30	49
Gambar 4.20 Metrik Evaluasi Regresi 80:20	50
Gambar 4.21 Hasil Evaluasi Model Klasifikasi Banjir 70:30	53
Gambar 4.22 Hasil Evaluasi Model Klasifikasi Banjir 80:20	55
Gambar 4.23 Hasil Prediksi 5 Hari	59
Gambar 4.24 Hasil Prediksi 5 Hari data Prediksi dan BMKG	60
Gambar 4.25 Grafik Prediksi 5 Hari.....	61

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	69
Lampiran 2. Curriculum Vitae	70
Lampiran 3. Surat Pernyataan & Pengalihan HAKI.....	71
Lampiran 4. Sertifikat BNSP	73
Lampiran 5. Surat Ijin Riset Perusahaan.....	74
Lampiran 6. Form Revisi Dosen Penguji.....	75

