

**KLASIFIKASI TINGKAT CURAH HUJAN HARIAN DI
STASIUN METEOROLOGI MARITIM TANJUNG
PRIOK DENGAN ALGORITMA *K-NEAREST
NEIGHBORS***

SKRIPSI

Oleh:

Muhammad Fadil

201910225170



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas akhir : Klasifikasi Tingkat Curah Hujan Harian Di Stasiun
Meteorologi Maritim Tanjung Priok Dengan Algoritma K-
Nearest Neighbors

Nama Mahasiswa : Muhammad Fadil

Nomor Pokok Mahasiswa : 201910225170

Program Studi/Fakultas : Informatika / Ilmu Komputer

Jakarta, 23 April 2025

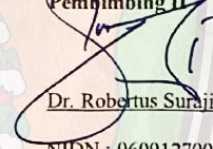
MENYETUJUI,

Pembimbing I

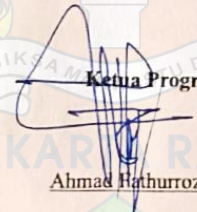

Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

Pembimbing II


Dr. Robertus Suraji, S.S., M.A.

NIDN : 0609127001


Ketua Program Studi

Ahmad Huthurrozi, S.E., MMSI.

NIP : 2012786

Program Studi Informatika

Fakultas Ilmu Komputer

Universitas Bhayangkara Jakarta Raya

2025

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul Tugas Akhir : Klasifikasi Tingkat Curah Hujan Harian
Di Stasiun Meteorologi Maritim Tanjung
Priok Dengan Algoritma *K-Nearest
Neighbors*

Nama Mahasiswa : Muhammad Fadil
Nomor Pokok Mahasiswa : 201910225170
Program Studi/Fakultas : Informatika/Ilmu Komputer
Tanggal Lulus Ujian Tugas Akhir : 25 Juni 2025

Jakarta, 07 Juli 2025

MENGESAHKAN,

Ketua Tim Penguji : Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107

Penguji I : Asep Ramdhani Mahbub, S.Kom., M.Kom.

NIDN : 0329087703

Penguji II : Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I.

NIP. 2012486

Dr. Robertus Suraji, S.S., M.A.

NIP. 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Muhammad Fadil
NPM : 201910225170
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Klasifikasi Tingkat Curah Hujan Harian di Stasiun
Meteorologi Maritim Tanjung Priok Dengan
Algoritma *K-Nearest Neighbors*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, April 2025

Penulis



Muhammad Fadil

ABSTRAK

Muhammad Fadil – 201910225170, Klasifikasi Tingkat Curah Hujan Harian Di Stasiun Meteorologi Maritim Tanjung Priok Dengan Algoritma *K-Nearest Neighbors*. Informatika Universitas Bhayangkara Jakarta Raya. 2025

Penelitian ini bertujuan untuk mengembangkan model klasifikasi intensitas curah hujan harian berdasarkan data historis dari Stasiun Meteorologi Maritim Tanjung Priok periode 2015–2023. Metode yang digunakan adalah algoritma *K-Nearest Neighbors* (KNN) dengan pendekatan CRISP-DM. Dataset terdiri dari 3.287 entri harian yang mencakup parameter cuaca seperti suhu, kelembapan, curah hujan, penyinaran matahari, dan kecepatan angin. Data praproses meliputi penanganan missing values, normalisasi fitur, pembentukan kategori curah hujan berdasarkan standar BMKG, serta penyeimbangan data menggunakan teknik *SMOTE*. Model dievaluasi menggunakan metode *10-fold cross-validation* dan optimasi parameter melalui *GridSearchCV*. Hasil penelitian menunjukkan bahwa model KNN dengan parameter terbaik ($k=7$, $weight='distance'$, $p=2$) mampu mengklasifikasikan curah hujan ke dalam lima kategori (tidak hujan hingga hujan sangat lebat) dengan akurasi rata-rata 90,38% dan standar deviasi 2,96%. Model ini berpotensi digunakan sebagai sistem pendukung keputusan dalam prediksi cuaca harian, khususnya di wilayah Jakarta Utara.

Kata Kunci: Curah Hujan, Klasifikasi, *K-Nearest Neighbors*, *CRISP-DM*, *SMOTE*, *Grid Search*, *Cross Validation*.

ABSTRACT

Muhammad Fadil – 201910225170, *Daily Rainfall Intensity Classification at the Maritime Meteorological Station of Tanjung Priok Using the K-Nearest Neighbors Algorithm*. Informatics, Bhayangkara University Jakarta Raya. 2025

This study aims to develop a classification model for daily rainfall intensity based on historical data from the Maritime Meteorology Station of Tanjung Priok for the period 2015–2023. The method used is the K-Nearest Neighbors (KNN) algorithm with the CRISP-DM approach. The dataset consists of 3,287 daily entries including weather parameters such as temperature, humidity, rainfall, solar radiation, and wind speed. The preprocessing stages include handling missing values, feature normalization, creating rainfall categories based on BMKG standards, and balancing the dataset using SMOTE. The model is evaluated using 10-fold cross-validation and parameter optimization via GridSearchCV. The results show that the best KNN model ($k=7$, $\text{weight}=\text{'distance'}$, $p=2$) can classify rainfall into five categories (no rain to very heavy rain) with an average accuracy of 90.38% and a standard deviation of 2.96%. This model has the potential to serve as a decision support system for daily weather prediction, particularly in North Jakarta.

Keywords: *Rainfall, Classification, K-Nearest Neighbors, CRISP-DM, SMOTE, Grid Search, Cross Validation.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Muhammad Fadil
NPM : 201910225170
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

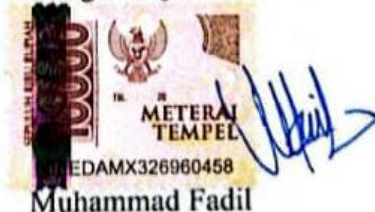
Klasifikasi Tingkat Curah Hujan Harian Di Stasiun Meteorologi Maritim Tanjung Priok Dengan Algoritma *K-Nearest Neighbors*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 8 Juli 2025
Yang Menyatakan


METERAI
TEMPEL
EDAMX326960458
Muhammad Fadil

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa, karena atas rahmat-Nya yang telah diberikan kepada penulis, sehingga dapat penulis dapat menyusun tugas akhir skripsi yang berjudul **“Klasifikasi Tingkat Curah Hujan Harian Di Stasiun Meteorologi Maritim Tanjung Priok Dengan Algoritma K-Nearest Neighbors”**, dengan baik dan tepat waktu sebagaimana dengan waktu yang telah ditentukan.

Dalam rangka penyusunan skripsi ini, dengan penuh rasa hormat dan rendah hati, penulis ingin menyampaikan penghargaan yang tulus kepada teristimewa kedua orang tua Ibu dan Bapa yang memberikan perhatian, dukungan, doa, dan semangat dalam menyelesaikan penelitian ini serta:

1. Bapak Irjen Pol. (Purn) Drs. H. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya dan sebagai Pembimbing II Skripsi.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Ibu Dr. Herlawati, S.Si., M.M., M.Kom. sebagai Dosen Pembimbing I Skripsi, penulis merasa sangat berterima kasih atas bimbingan dan motivasi yang telah diberikan dalam proses penyusunan skripsi.

5. Ibu Dwi Budi Srisulistiowati, S.Kom., M.M. selaku dosen pembimbing akademik.
6. Jasmine kinari, Mia Alvianita, Veronika Valentina, Inesyah Irena Zaelsyah Nassir, Bagus Prasetyo, Juandika Shevani, Fakhri Ramadan, Ahmad Firmansyah, Al Anwari Abbas, Muhamad Arshi, Winner Siagian, Bangun Tegar dan teman teman A4, yang telah banyak membantu saya dan selalu memberikan saya banyak dukungan dan kasih sayang baik secara mental dan fisik.

Penulis menyadari bahwa dalam penyusunan tugas akhir skripsi ini terdapat banyak kekurangan dan masih jauh dari kesempurnaan. Oleh karena itu, penulis dengan tulus menerima kritik dan saran yang membangun dari pembaca, dengan harapan dapat meningkatkan penelitian ini menjadi lebih baik. Penulis berharap agar penelitian ini dapat memberikan manfaat bagi pembaca secara umum, terutama bagi penulis sendiri sebagai penyusun, seluruh mahasiswa Universitas Bhayangkara Jakarta Raya, dan perusahaan terkait.

Jakarta, 30 Desember 2024



Muhammad Fadil

201910225170

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN.....	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI.....	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Batasan Masalah.....	6

1.7 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terkait	9
2.2 <i>Machine Learning</i>	11
2.3 <i>Data Mining</i>	13
2.4 <i>K-Nearest Neighbors</i>	15
2.5 <i>CRIPS-DM</i>	18
2.6 <i>Cross Validation</i>	20
2.7 <i>Grid Search</i>	21
2.8 Curah Hujan	22
2.9 Klasifikasi	23
2.10 <i>Google Colab</i>	24
BAB III METODOLOGI PENELITIAN.....	26
3.1 Objek Penelitian	26
3.2 Waktu Penelitian	27
3.3 Kerangka Penelitian.....	27
3.4 Metode Pengumpulan Data.....	34
3.5 Metode Analisis Data	36

BAB IV HASIL DAN PEMBAHASAN	41
4.1 <i>Business Understanding</i> (Pemahaman Bisnis).....	41
4.2 <i>Data Understanding</i> (Pemahaman Data).....	42
4.2.1 <i>Membaca Dataset</i>	42
4.3 <i>Data Preparation</i> (Persiapan Data).....	43
4.3.1 Penanganan Missing Values	43
4.3.2 Normalisasi Data.....	44
4.3.3 Membuat Kategori Hujan.....	45
4.3.4 Pemilihan Variabel.....	46
4.3.5 Penanganan Ketidakseimbangan Data	47
4.3.6 Penyimpanan Dataset Akhir Yang Telah Dibersihkan	48
4.4 <i>Modeling</i> (Pemodelan).....	49
4.4.1 <i>Split Data</i> (Pembagian data latih dan data uji)	49
4.4.2 Pemodelan KNN	49
4.5 <i>Evaluation</i> (Evaluasi)	51
4.5.1 Hasil Prediksi dan Evaluasi Akurasi	51
4.5.2 Validasi Silang (<i>Cross-Validation</i>).....	52
4.5.3 Visualisasi <i>Cross-Validation Scores</i>	53

4.6 <i>Deployment</i>	54
BAB V PENUTUP	55
5.1 Kesimpulan	55
5.2 Saran	56
DAFTAR PUSTAKA	59
LAMPIRAN	63



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	9
Tabel 3.1 Tabel implementasi	27



DAFTAR GAMBAR

Gambar 2. 1 Machine Learning[13]	12
Gambar 2. 2 Data Mining[16]	14
Gambar 2. 3 Ilustrasi KNN[13]	15
Gambar 2. 4 CRISP-DM[18].....	18
Gambar 2. 5 Google Colab.....	25
Gambar 3. 1 Peta administratif Kota Jakarta Utara	26
Gambar 3. 2 Kerangka Penelitian.....	28
Gambar 3. 3 Halaman login aplikasi Data Online BMKG	30
Gambar 3. 4 Tampilan beranda setelah berhasil masuk.....	30
Gambar 3. 5 From pencarian data harian.....	31
Gambar 3. 6 Hasil pencarian dan opsi unduh data stasiun dan parameter.....	32
Gambar 3. 7 Algoritma K-Nearest Neighbors.....	33
Gambar 3. 8 Dataset.....	34
Gambar 3. 9 Diagram Penelitian Algoritma KNN	40
Gambar 4. 1 Membaca Dataset.....	42
Gambar 4. 2 Penanganan Missing Values.....	43
Gambar 4. 3 Impute Missing Values	44
Gambar 4. 4 Pengecekan Ulang Missing Values	44
Gambar 4. 5 Normalisasi Data	45
Gambar 4. 6 Membuat Kategori dan Kolom Kategori	46

Gambar 4. 7 Pemilihan Fitur	47
Gambar 4. 8 Penanganan ketidakseimbangan kelas	48
Gambar 4. 9 Penyimpanan Data Bersih	48
Gambar 4. 10 Split Data.....	49
Gambar 4. 11 Modeling	50
Gambar 4. 12 Evaluasi Akurasi.....	51
Gambar 4. 13 Validasi Silang (Cross-Validation).....	52
Gambar 4. 14 Visualisasi Cross-Validation Scores.....	53
Gambar 4.15 Deployment	54



DAFTAR LAMPIRAN

Lampiran 1 Hasil Plagiarisme	64
Lampiran 2 Biodata Mahasiswa	65
Lampiran 3 Kartu Bimbingan Skripsi	66

