

**KLASIFIKASI CURAH HUJAN HARIAN MENGGUNAKAN
METODE K-NEAREST NEIGHBOR DENGAN OPTIMASI
ALGORITMA GENETIKA**

SKRIPSI



Oleh :

MUHAMMAD ALVIRIZA RAMADHAN

NPM. 20081010143

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2024**

LEMBAR PENGESAHAN SKRIPSI

**Judul : KLAIFIKASI CURAH HUJAN HARIAN MENGGUNAKAN
METODE K-NEAREST NEIGHBOR DENGAN OPTIMASI
ALGORITMA GENETIKA**

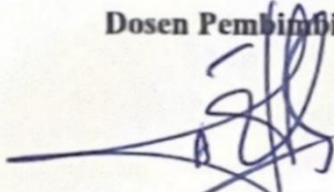
Oleh : Muhammad Alviriza Ramadhan

NPM : 20081010143

**Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Selasa, Tanggal 21 Mei 2024**

Mengetahui

1. **Dosen Pembimbing**



Fetty Tri Anggraeny, S.Kom., M.Kom.
NIP. 19820211 2021212 005

1. **Dosen Penguji**



Retno Mumpuni, S.Kom., M.Sc
NPT. 172198 70 716054

2.



Christia Wji Putra, S.Kom., M.T
NIP. 19861008 2021211 001

2.



Faywaz Ali Akbar., S.Kom., M.Kom.
NIP. 19920317 2018031 002

Menyetujui

**Dekan
Fakultas Ilmu Komputer**



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

**Koordinator Program Studi
Informatika**



Fetty Tri Anggraeny, S.Kom., M.Kom.
NIP. 19820211 2021212 005

SURAT KETERANGAN ANTI PLAGIAT

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Muhammad Alviriza Ramadhan

NPM : 20081010143

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

“KLASIFIKASI CURAH HUJAN HARIAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN OPTIMASI ALGORITMA GENETIKA”

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dari juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.

Surabaya, 21 Mei 2024

Hormat saya,



MUHAMMAD ALVIRIZA R.
NPM. 20081010143

Nama Mahasiswa : Muhammad Alviriza Ramadhan

NPM : 20081010143

Program Studi : Informatika

Dosen Pembimbing : Fetty Tri Anggraini, S.Kom, M.Kom

Chrystia Aji Puta, S.Kom, M.T

ABSTRAK

Cuaca merupakan kondisi udara yang diamati dalam periode yang relatif singkat atau pendek pada teritori yang terbatas. Kemajuan teknologi pada saat ini sangat membantu manusia khususnya dalam hal pengamatan cuaca seperti halnya pengklasifikasian curah hujan dalam skala harian. Salah satu metode yang digunakan yaitu penggunaan metode K-NN serta algoritma genetika. Perubahan cuaca dan iklim dapat dipengaruhi oleh banyak faktor seperti adalah suhu, kelembaban, tekanan udara, kecepatan angin, arah angin saat kecepatan maksimum, lama penyinaran matahari, dan curah hujan. Maka dilakukannya Klasifikasi curah hujan harian menggunakan metode KNN dengan optimasi algoritma genetika” dengan output yang diharapkan yaitu tingkat akurasi dari penerapan metode yang digunakan. Hasil dari penelitian ini adalah Dataset yang dilatih dan diuji menggunakan rasio pembagian data training dan data testing sebesar 60:40 cenderung menghasilkan Tingkat akurasi yang lebih tinggi. Metode K-Nearest Neighbor dengan optimasi algoritma genetika yang dieksekusi menggunakan parameter jumlah generasi sebanyak 50 cenderung menghasilkan Tingkat akurasi yang lebih tinggi. Metode K-Nearest Neighbor dengan optimasi algoritma genetika yang dieksekusi menggunakan parameter probabilitas crossover sebesar 0,5 ketika dikombinasikan dengan parameter jumlah generasi sebanyak 50 menghasilkan Tingkat akurasi yang lebih besar dibandingkan dengan menggunakan parameter probabilitas crossover sebesar 0,75 yang dikombinasikan dengan parameter jumlah generasi yang bernilai sama.

Kata kunci: Curah Hujan, KNN, Algoritma Genetika

KATA PENGANTAR

Puji Syukur kehadiran Allah SWT yang telah memberikan keberkahan dan kesehatan kepada penulis, sehingga penulis dapat menyelesaikan kegiatan penelitian skripsi ini dengan judul

“KLASIFIKASI CURAH HUJAN HARIAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN OPTIMASI ALGORITMA GENETIKA”

Skripsi ini dibuat penulis untuk memenuhi mata kuliah skripsi sebagai syarat untuk menyelesaikan Strata Satu (S1) dari Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penulis berharap dengan adanya penyusunan skripsi ini dapat menambah ilmu baru dan wawasan yang bermanfaat bagi semua pihak yang membaca.

Penulis menyadari dalam penulisan laporan penelitian skripsi ini masih jauh dari kata sempurna. Atas segala kekurangan dan ketidaksempurnaan laporan skripsi ini, maka penulis sangat mengharapkan kritik, saran, dan masukan yang bersifat membangun kearah perbaikan dan penyempurnaan.

Surabaya, 21 Mei 2024

Muhammad Alviriza R.

UCAPAN TERIMA KASIH

Ucapan terima kasih ini penulis persembahkan sebagai perwujudan rasa syukur atas terselesaikannya Laporan Skripsi ini. Dalam selesainya Laporan Skripsi ini, penulis menyadari bahwa tidak akan berhasil tanpa adanya bantuan, dukungan, serta bimbingan yang telah memberikan masukan-masukan, kritik dan saran dari berbagai pihak. Oleh karena itu pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT., IPU. selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur serta selaku Dosen Pembimbing satu yang telah memberikan banyak masukan serta membantu dengan sabar dalam membimbing penulis untuk pengerjaan skripsi serta jurnal penulis dengan maksimal.
4. Ibu Made Hanindia Prameswari, S.Kom., M.Cs. selaku Dosen Wali yang telah banyak membantu penulis dari awal perkuliahan hingga terselesaikannya skripsi.
5. Bapak Chrystia Aji Putra, S.Kom., MT. selaku Dosen Pembimbing dua yang telah membantu penulis dengan sabar dalam proses pengerjaan skripsi penulis secara maksimal.
6. Bapak Andreas Nugroho Sihananto, S.Kom., M.Kom. selaku koordinator skripsi jurusan Informatika yang membantu proses administrasi dan keberlangsungan sidang Skripsi.
7. Seluruh Dosen Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur atas segala ilmu pengetahuan yang diberikan kepada penulis selama masa perkuliahan. Semoga penulis dapat mengamalkan ilmu pengetahuan yang telah diberikan dan akan menjadi bekal yang berguna bagi penulis di masa yang akan datang.

8. Papa dan Mama yang telah memberikan dukungan secara materi dan non – materi sekaligus selalu mendoakan penulis dalam kondisi apapun. Sehingga penulis dapat menyelesaikan perkuliahan dari awal hingga terselesaikannya skripsi ini dengan baik.
9. Nur Racana Kuslaila yang telah membantu penulis dengan meminjamkan laptop untuk proses pengerjaan skripsi ini dari awal hingga skripsi ini terselesaikan serta memberikan bantuan secara moral kepada penulis dari awal bertemu hingga saat ini.
10. Arif Widiawan Subagio, Fahmi Anugrah Danendra, Ahmad Hasby Bik, Rama Ariya Candra yang telah membantu memberikan arahan kepada penulis dalam proses pengerjaan program dan laporan skripsi dengan baik.
11. Himpunan Mahasiswa Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur (HIMATIFA UPN “Veteran” Jawa Timur) dan Badan Legislatif Mahasiswa Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur (BLM FASILKOM UPN “Veteran” Jawa Timur) yang telah menjadi wadah bagi penulis untuk belajar berorganisasi, menambah relasi serta melatih manajemen yang dapat berguna bagi masa depan penulis.
12. Seluruh teman – teman angkatan 2020 Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur yang tidak dapat penulis sebutkan satu per satu yang telah berjuang bersama dalam proses perkuliahan dari awal hingga selesai.

Semoga Tuhan Yang Maha Esa memberikan balasan yang setimpal kepada semua pihak yang telah memberikan bantuan, bimbingan maupun motivasi yang bermanfaat bagi penulis. Penulis juga menyadari bahwa masih banyak kekurangan pada penulisan Laporan Skripsi ini. Oleh karena itu, saran dan kritik yang membangun akan penulis terima dengan senang hati dan penulis berharap semoga Laporan Skripsi ini dapat memberikan manfaat dalam perkembangan ilmu pengetahuan bagi pembaca.

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
SURAT KETERANGAN ANTI PLAGIAT.....	ii
ABSTRAK	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH.....	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan	4
1.4. Manfaat	4
1.5. Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian Pendahulu.....	6
2.2. Perbandingan Penelitian Pendahulu.....	13
2.3. Cuaca	14
2.3.1. Suhu Udara.....	15
2.3.2. Tekanan Udara	15
2.3.3. Kecepatan Angin.....	15
2.3.4. Kelembaban Udara.....	15
2.4. Curah Hujan	15
2.5. Deret Waktu (<i>Time Series</i>).....	16
2.6. Pra Proses Data	16

2.6.1. Data Cleansing (Pembersihan Data)	17
2.6.2. Data Editing (Pengeditan Data)	17
2.6.3. Data Reduction (Pengurangan Data).....	18
2.6.4. Data Wrangling (Penggulungan Data).....	18
2.7. Python	19
2.8. Google Collaboratory	19
2.9. Algoritma Genetika.....	19
2.10. Komponen Algoritma Genetika	21
2.10.1. Teknik Pengkodean.....	21
2.10.2. Prosedur Inisialisasi (generate populasi awal).....	21
2.10.3. Fungsi Evaluasi	22
2.10.4. Seleksi	23
2.10.5. Operator Genetika	23
2.11. Parameter Kontrol	24
2.11.1. Probabilitas Crossover (Pc).....	25
2.11.2. Probabilitas Mutasi (Pm)	25
2.12. Metode K-Nearest Neighbor.....	25
BAB III METODOLOGI PENELITIAN.....	27
3.1. Analisis Kebutuhan	27
3.1.1. Lingkungan Penelitian	27
3.1.2. Spesifikasi Perangkat Keras.....	27
3.1.3. Spesifikasi Perangkat Lunak	28
3.2. Tahapan Penelitian	28
3.2.1. Identifikasi Permasalahan	29
3.2.2. Studi Literatur	29

3.2.3. Pembuatan Program	30
3.2.4. Menyimpulkan Hasil	30
3.3. Perancangan Sistem	30
3.4. Pengumpulan Data	32
3.5. Pre-processing Data	33
3.5.1. Dataset	33
3.5.2. Missing Value	33
3.5.3. Label Encoding	34
3.5.4. Split Data	34
3.6. Split Data Training dan Data Testing	35
3.7. Membangkitkan Populasi Awal	35
3.8. Evaluasi Fitness Menggunakan KNN	36
3.8.1. Input Parameter K	38
3.8.2. Menghitung Jarak	38
3.8.3. Mengurutkan Jarak	38
3.8.4. Mengumpulkan Kategori	38
3.8.5. Mencari Kategori yang Sering Muncul	38
3.8.6. Output Nilai Fitness Tiap Individu	38
3.9. Seleksi Turnamen	39
3.10. Crossover Two Point	40
3.11. Mutasi	41
3.12. Individu Baru	41
BAB IV HASIL DAN PEMBAHASAN	42
4.1. Penyiapan Dataset	42
4.2. Analisa Dataset	42

4.3. Pre-processing Data	44
4.3.1. Data Selection	44
4.3.2. Data Cleaning	46
4.3.3. Pelabelan Kelas	52
4.4. Data Splitting	53
4.4.1. Pembagian Dengan Rasio 60:40	53
4.4.2. Pembagian Dengan Rasio 70:30	54
4.5. Implementasi K-Nearest Neighbor Dengan Optimasi Algoritma Genetika	55
4.5.1. Menentukan Populasi Awal	55
4.5.2. Evaluasi Fitness	56
4.5.3. Seleksi Turnamen	57
4.5.4. Crossover	59
4.5.5. Mutasi	60
4.5.6. Pergantian Generasi	61
4.5.7. Evaluasi Hasil	63
4.6. Implementasi Program	64
4.7. Skenario Uji Coba	70
4.7.1. Uji Coba Pertama	70
4.7.2. Uji Coba Kedua	71
4.7.3. Uji Coba Ketiga	71
4.7.4. Uji Coba Keempat	72
4.7.5. Uji Coba Kelima	72
4.7.6. Uji Coba Keenam	73
4.7.7. Uji Coba Ketujuh	73

4.7.8. Uji Coba Kedelapan	74
4.7.9. Uji Coba Kesembilan	75
4.7.10. Uji Coba Kesepuluh	75
4.7.11. Uji Coba Kesebelas	76
4.7.12. Uji Coba Keduabelas.....	76
4.7.13. Uji Coba Ketigabelas	77
4.7.14. Uji Coba Keempatbelas.....	77
4.7.15. Uji Coba Kelimabelas	78
4.7.16. Uji Coba Keenambelas.....	78
BAB V KESIMPULAN DAN SARAN.....	80
5.1. Kesimpulan	80
5.2. Saran	81
DAFTAR PUSTAKA	82
BIODATA PENULIS	84

DAFTAR TABEL

Tabel 2.2.1 Tabel Perbandingan Penelitian Pendahulu	13
Tabel 3.1.1 Spesifikasi Perangkat Keras	28
Tabel 3.1.2 Spesifikasi Perangkat Lunak	28
Tabel 3.4.1 Dataset Yang Digunakan.....	32
Tabel 3.5.1 Dataset Setelah Melalui Proses Pre-Processing	34
Tabel 4.2.1 Tabel Data Awal.....	43
Tabel 4.3.1 Tabel Variabel Yang Digunakan.....	44
Tabel 4.3.2 Tabel Data Yang Digunakan	45
Tabel 4.3.3 Tabel Perubahan Nama Variabel Data	46
Tabel 4.3.4 Potongan Kode 1	47
Tabel 4.3.5 Potongan Kode 2	48
Tabel 4.3.6 Potongan Kode 3	49
Tabel 4.3.7 Potongan Kode 4	51
Tabel 4.3.8 Tabel Ringkasan Akumulasi Data.....	51
Tabel 4.3.9 Tabel Ringkasan Statistik Data	52
Tabel 4.3.10 Kelas Curah Hujan Harian Menurut BMKG.....	52
Tabel 4.3.11 Tabel Akumulasi Jumlah Data Pada Setiap Kelas	53
Tabel 4.5.1 Kode Program 5	56
Tabel 4.5.2 Potongan Kode 6	57

Tabel 4.5.3 Potongan Kode 7	58
Tabel 4.5.4 Potongan Kode 8	59
Tabel 4.5.5 Potongan Kode 9	61
Tabel 4.5.6 Potongan Kode 10	62
Tabel 4.5.7 Potongan Kode 11	63
Tabel 4.6.1 Potongan Kode 12	64
Tabel 4.6.2 Potongan Kode 13	65
Tabel 4.6.3 Potongan Kode 14	65
Tabel 4.6.4 Potongan Kode 15	66
Tabel 4.6.5 Potongan Kode 16	66
Tabel 4.6.6 Potongan Kode 17	67
Tabel 4.6.7 Potongan Kode 18	68
Tabel 4.7.1 Ringkasan Data Skenario Pengujian	79

DAFTAR GAMBAR

Gambar 3.2.1 Diagram Alur Tahapan Penelitian	29
Gambar 3.3.1 Perancangan Sistem	31
Gambar 3.5.1 Tahapan Pre-Processing Data	33
Gambar 3.8.1 Algoritma KNN	37
Gambar 3.9.1 Ilustrasi Seleksi Turnamen	40
Gambar 3.10.1 Contoh Dua Individu	40
Gambar 3.10.2 Hasil Crossover Dua Titik	41
Gambar 3.11.1 Ilustrasi Mutasi	41
Gambar 4.3.1 Output Potongan Kode 1	48
Gambar 4.3.2 Output Potongan Kode 2	49
Gambar 4.3.3 Output Potongan Kode 3	50
Gambar 4.4.1 Akumulasi Data Perkelas Pada Rasio 60:40.....	54
Gambar 4.4.2 Akumulasi Data Perkelas Pada Rasio 70:30.....	54
Gambar 4.6.1 Statistik Penginstalan Library DEAP	65
Gambar 4.6.2 Output Dari Metode Knn Dengan Optimasi Algoritma Genetika	69
Gambar 4.6.3 Grafik Dari Metode Knn Dengan Optimasi Algoritma Genetika.	69
Gambar 4.7.1 Hasil Uji Coba Pertama	71
Gambar 4.7.2 Hasil Uji Coba Kedua.....	71
Gambar 4.7.3 Hasil Uji Coba Ketiga.....	72

Gambar 4.7.4 Hasil Uji Coba Keempat	72
Gambar 4.7.5 Hasil Uji Coba Kelima.....	73
Gambar 4.7.6 Hasil Uji Coba Keenam	73
Gambar 4.7.7 Hasil Uji Coba Ketujuh	74
Gambar 4.7.8 Hasil Uji Coba Kedelapan	74
Gambar 4.7.9 Hasil Uji Coba Kesembilan	75
Gambar 4.7.10 Hasil Uji Coba Kesepuluh	75
Gambar 4.7.11 Hasil Uji Coba Kesebelas	76
Gambar 4.7.12 Hasil Uji Coba Keduabelas.....	76
Gambar 4.7.13 Hasil Uji Coba Ketigabelas	77
Gambar 4.7.14 Hasil Uji Coba Keempatbelas.....	77
Gambar 4.7.15 Hasil Uji Coba Kelimabelas	78
Gambar 4.7.16 Hasil Uji Coba Keenambelas.....	78